Cochrane Library

Cochrane Database of Systematic Reviews

Bioidentical hormones for women with vasomotor symptoms (Review)

Gaudard AMIS, Silva de Souza S, Puga MES, Marjoribanks J, da Silva EMK, Torloni MR

Gaudard AMIS, Silva de Souza S, Puga MES, Marjoribanks J, da Silva EMK, Torloni MR.
Bioidentical hormones for women with vasomotor symptoms.
Cochrane Database of Systematic Reviews 2016, Issue 8. Art. No.: CD010407.
DOI: 10.1002/14651858.CD010407.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Bioidentical hormones for women with vasomotor symptoms

Ana Marcia IS Gaudard[1], Sulani Silva de Souza[1], Maria ES Puga[2], Jane Marjoribanks[3], Edina MK da Silva[4], Maria R Torloni[2]

[1]Management of educational evaluation, School of Sciences of Health/FEPECS, Brasília, Brazil. [2]Brazilian Cochrane Centre, Centro de Estudos de Saúde Baseada em Evidências e Avaliação Tecnológica em Saúde, São Paulo, Brazil. [3]Department of Obstetrics and Gynaecology, University of Auckland, Auckland, New Zealand. [4]Emergency Medicine and Evidence Based Medicine, Universidade Federal de São Paulo, São Paulo, Brazil

Contact address: Ana Marcia IS Gaudard, Management of educational evaluation, School of Sciences of Health/FEPECS, SMHN Quadra 3 conjunto A Bloco 1 Edifício FEPECS, Brasília, 70710-100, Brazil. marciayunesg@uol.com.br.

Editorial group: Cochrane Gynaecology and Fertility Group.
Publication status and date: New, published in Issue 8, 2016.

Citation: Gaudard AMIS, Silva de Souza S, Puga MES, Marjoribanks J, da Silva EMK, Torloni MR. Bioidentical hormones for women with vasomotor symptoms. *Cochrane Database of Systematic Reviews* 2016, Issue 8. Art. No.: CD010407. DOI: 10.1002/14651858.CD010407.pub2.

ABSTRACT

Background

Various hormone therapies (HT) are available to treat menopausal vasomotor symptoms. Bioidentical hormones are chemically identical to those produced by the human body, and several types are well-tested and available on prescription. Many women have opted for bioidentical hormone therapy (BHT) on the assumption that it is safer than other forms of HT. We evaluated the evidence.

Objectives

To determine the effectiveness and safety of bioidentical hormones compared to placebo or non-bioidentical hormones for the relief of vasomotor symptoms.

Search methods

In July 2015 we searched the Cochrane Central Register of Controlled Trials, PubMed, Embase, Literatura Latino-Americana e do Caribe em Ciências da Saúde (LILACS), registers of ongoing trials and the reference lists of articles retrieved.

Selection criteria

Randomised controlled trials (RCTs) comparing bioidentical hormone therapy (BHT) versus placebo or non-bioidentical hormones.

Data collection and analysis

We used standard methodological procedures expected by the Cochrane Collaboration. Our primary outcome was vasomotor symptoms (hot flushes and night sweats). We evaluated the overall quality of the evidence using Grading of Recommendations Assessment, Development and Evaluation criteria (GRADE).

Main results

We included 23 RCTs (5779 participants). Most studies (20/23) included only women with moderate to severe hot flushes. All studies compared unopposed 17 beta-estradiol (beta-estradiol) versus placebo or conjugated equine estrogens (CEE). None of the studies reported night sweats as a separate outcome.

BHT patch versus placebo

Frequency of hot flushes

Four RCTs reported data suitable for analysis. There were fewer hot flushes in the BHT group, with a moderate to large effect size (SMD -0.68, 95% CI -0.83 to -0.53, four RCTs, 793 women, I^2 = 67%, low quality evidence). There was moderate heterogeneity, but a consistent direction of effect. Seven RCTs reported data unsuitable for analysis; all reported a benefit in the intervention group.

Symptom intensity

Two RCTs reported analysable data. Measured on a 0-100 visual analogue scale (VAS), hot flush intensity was lower in the BHT group (MD -19.94 points, 95% CI -24.86 to -15.02, two RCTs, 393 women, I^2 = 54%, low quality evidence). There was moderate heterogeneity, but a consistent direction of effect.

Adverse effects

Adverse events (such as headache, vaginal bleeding, breast tenderness and skin reactions) were more common in the intervention group (odds ratio (OR) 2.14, 95% CI 1.29 to 3.54, 9 RCTs, 1822 women, I^2 = 73%, low quality evidence). There was moderate heterogeneity, but a consistent direction of effect. In one study, five women in the intervention group developed endometrial hyperplasia.

BHT gel versus placebo

Hot flush frequency

Three RCTs reported this outcome, but the data were unsuitable for analysis. All reported a benefit in the BHT group.

Adverse effects

Adverse events were more common in the BHT group (OR 1.41, 95% CI 1.09 to 1.83, 3 RCTs, 1086 women, I^2 = 0%, moderate quality evidence).

Oral BHT versus placebo

Hot flush frequency

Two studies reported analysable data. There were fewer hot flushes in the BHT group, with a moderate to large effect size (SMD -0.80, 95% CI -1.03 to -0.57, two RCTs, 356 women, I^2 = 14%, low quality evidence).

Adverse effects

There was no evidence of a difference between the groups (OR 1.28, 95% CI 0.84 to 1.96, 3 RCTs, 433 women, I^2 = 0%, low quality evidence).

Topical BHT emulsion versus placebo

Hot flush frequency

One study with data unsuitable for analysis reported a benefit in the intervention group.

Adverse effects

There was no evidence of a difference between the groups (OR 1.46, 95% CI 0.80 to 2.66, one RCT, 200 women, low quality evidence).

Intranasal BHT versus placebo

Hot flush frequency

Only one study reported analysable data. There were fewer hot flushes per day in the BHT group (MD -3.04 95% CI -4.05 to -2.03, one study, 458 women, moderate quality evidence)

Adverse effects

Adverse events (such as headache, breast tenderness, arthralgia and nausea) were more common in the intervention group (OR 1.96, 95% CI 1.26 to 3.03, one RCT, 458 women, moderate quality evidence).

Subgroup analyses

Subgroup analyses by dose of BHT suggested that higher doses of BHT may be associated with more effectiveness but also higher risk of adverse effects.

BHT patch versus 0.625 mg CEE

Two RCTs reported this comparison, but the data were unsuitable for analysis.

Hot flush frequency

Both RCTs reported no evidence of a difference between the groups.

Adverse effects

Findings were inconsistent. In one comparison (0.1 mg BHT versus CEE), breast pain and vaginal bleeding were more frequent in the BHT group.

Oral BHT versus 0.625 mg CEE

Hot flush frequency

One study with data unsuitable for analysis reported no evidence of a difference between the groups.

Adverse effects

There was no evidence of a difference between the groups (OR 1.20, 95% CI 0.50 to 2.87, one RCT, 103 women, very low quality evidence).

Authors' conclusions

There was low to moderate quality evidence that BHT in various forms and doses is more effective than placebo for treating moderate to severe menopausal hot flushes. There was low to moderate quality evidence of higher rates of adverse effects such as headache, vaginal bleeding, breast tenderness and skin reactions in the BHT group. There was some evidence to suggest that higher doses of BHT are associated with greater effectiveness but also with higher risk of adverse effects. Although all the included studies used unopposed estrogen, it is recommended best practice to use progestogen therapy in women with a uterus taking estrogen in order to avoid endometrial hyperplasia, regardless of the source of the estrogen. No data are yet available about the safety of BHT with regard to long-term outcomes such as heart attack, stroke and breast cancer.

There was no good evidence of a difference in effectiveness between BHT and CEE, and findings with regard to adverse effects were inconsistent. The quality of the evidence was too low to reach any firm conclusions.

The main limitations in the quality of the evidence were study risk of bias (mainly due to poor reporting of methods), imprecision and lack of data suitable for analysis.

PLAIN LANGUAGE SUMMARY

Bioidentical hormones for vasomotor menopausal symptoms (hot flushes or night sweats)

Review question

This Cochrane review evaluates the effectiveness and safety of bioidentical hormone treatment (BHT) compared to no treatment or non-bioidentical hormone treatment (HT) for vasomotor symptoms experienced during the menopausal transition period.

Background

Various hormone therapies (HT) are available to treat menopausal vasomotor symptoms. Bioidentical hormones are chemically identical to those produced by the human body, and several types are well-tested and available on prescription. Many women have opted for bioidentical hormone therapy (BHT) on the assumption that it would be safer than other forms of HT. However, as it is unclear whether BHT is better or safer than other forms of HT, we evaluated the evidence.

Study characteristics

This review includes 23 randomised controlled trials conducted up to July 2015. These studies included a total of 5779 women who were in the menopausal transition period and suffered from hot flushes. Most of the studies (20/23) included only women with moderate to severe hot flushes. None of the studies reported night sweats as a separate outcome.

Key results

There is low to moderate quality evidence that BHT in various forms and doses is more effective than placebo in decreasing the frequency of moderate to severe hot flushes in women in the menopausal transition period. There was low to moderate quality evidence of higher rates of adverse effects such as headache, vaginal bleeding, breast tenderness and skin reactions in the BHT group. There is some evidence to suggest that higher doses of BHT are associated with more effectiveness but also higher risk of adverse effects. No data are yet available about the safety of BHT with regard to long-term outcomes such as heart attack, stroke and breast cancer. All women with a uterus who are taking any form of estrogen require co-administration of a progestogen, as unopposed estrogen is associated with endometrial hyperplasia.

There is no good evidence of a difference in effectiveness between BHT and CEE, and findings with regard to adverse effects are inconsistent. The quality of the evidence is too low to reach any firm conclusions for this comparison.

Quality of the evidence

The main limitations in the quality of the evidence were study risk of bias (mainly due to poor reporting of methods), imprecision and lack of data suitable for analysis.

SUMMARY OF FINDINGS FOR THE MAIN COMPARISON [Explanation]

Transdermal beta-estradiol patch versus placebo for women with hot flushes

Population: women with hot flushes
Setting: community
Intervention: beta-estradiol transdermal patch
Comparison: placebo

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Transdermal beta-estradiol patch versus placebo for women with hot flushes					
Frequency of hot flushes Beta-estradiol patch 0.0375-0.10 mg/day	There were fewer hot flushes in the beta-estradiol group. The effect size was moderate to large (SMD -0.68, 95% CI -0.83 to -0.53)			793 (4 RCTs)	⊕⊕◯◯ **low**[1,2]	Eight studies with data unsuitable for analysis all found a benefit for the beta-estradiol group. They utilised doses ranging from 0.014 mg/d to 2 mg/d
Intensity of hot flushes Beta-estradiol patch 0.025-0.05 mg/day	Measured on a 0-100 VAS, the intensity of hot flushes was lower in the beta-estradiol group (MD -19.94 points, 95% CI -24.86 to -15.02)			393 (2 RCTs)	⊕⊕◯◯ **low**[1,3,4]	
Adverse effects Beta-estradiol patch dose 0.10 mg/day	Rate in placebo group: 144 per 1000	Rate in beta-estradiol group*: 265 per 1000 (178 to 373)	OR 2.14 (1.29 to 3.54)	1822 (9 RCTs)	⊕⊕◯◯ low[1,5]	The rate of adverse effects was higher in the beta-estradiol group

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the median risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval; **OR:** Odds ratio; MD: mean difference; SMD: standardised mean difference; mg/d: milligrams per day

GRADE Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

1. Downgraded one level for serious risk of bias: inadequate explanation of study methods
2. Downgraded one level for serious inconsistency: I^2= 67%. Heterogeneity due to unexplained larger effect in the lowest-dose study. Direction of effect consistent
3. Downgraded one level for serious inconsistency: I^2= 54%. Heterogeneity due to larger effect in the highest-dose study. Direction of effect consistent
4. Downgraded one level for serious imprecision: small overall sample size
5. Downgraded one level for serious unexplained inconsistency: I^2= 73%. Direction of effect consistent

BACKGROUND

Description of the condition

According to the criteria proposed by the Stages of Reproductive Aging Workshop (STRAW), the late menopausal transition (stage 1) is marked by the occurrence of times when menstrual periods are absent (amenorrhoea) for periods lasting 60 days or more. The late menopausal transition is characterised by increased variability in menstrual cycle length, extreme fluctuations in hormonal levels, and increased prevalence of anovulation (failure of a woman's ovary to produce eggs). Follicle stimulating hormone (FSH) levels are sometimes elevated into the menopausal range (over 25 IU/L) and sometimes they are within the range of the earlier reproductive years, especially in association with high levels of 17 beta-estradiol (beta-estradiol). This stage is estimated to last an average of one to three years. Symptoms, most notably vasomotor symptoms, are likely to occur during this stage (Harlow 2012) which usually occurs between 40 to 65 years of age and corresponds to the period when a woman passes from the reproductive stage of life through the premenopausal transition and menopause to the postmenopausal years (Speroff 2011; Sturdee 2011). The hormone change associated with the menopausal transition can lead to a wide variety of symptoms that may negatively affect a woman's quality of life. The most common symptoms include hot flushes (flashes), night sweats, emotional lability, poor concentration and sleep disturbances. These symptoms can range from mild to severe (Speroff 2011).

The most common symptoms associated with menopause are hot flushes, night sweats, sleep disturbance, vaginal atrophy, and dyspareunia (NAMS 2012) In most women, these symptoms persist for a year or two after the menopause, but in some they may continue for 10 or more years. Vasomotor symptoms include hot flushes and night sweats. The hot flush is described as a sudden onset of reddening of the skin over the head, neck and chest accompanied by a feeling of warmth, often associated with spontaneous sweating, palpitations, and anxiety, resulting from a vasomotor response caused by decreased beta-estradiol levels (Nelson 2006). This aura is followed by measurable increased heat over the entire body surface. The duration of these episodes varies from a few seconds to several minutes and rarely they can last up to one hour. Flushes are more frequent and severe at night (when they are called night sweats) or during times of stress (Kronnenberg 1992).

The physiology of vasomotor symptoms is still not yet completely understood, but they apparently originate in the hypothalamus and are brought about by a decline in beta-estradiol. However, not all hot flushes are due to beta-estradiol deficiency. The correlation between the onset of flushes and beta-estradiol reduction is clinically supported by the effectiveness of beta-estradiol therapy and absence of flushes in permanent hypo-estradiol states, such as gonadal dysgenesis (Freedman 2001; Freedman 2006; Wilkin 1981). Among the many theories to explain why hot flushes occur, one of the most accepted (Tartaryn 1981) hypothesises that hot flushes are thermoregulatory events that aim to keep the body´ s temperature within a narrow thermoneutral zone. The brain has a thermoregulating centre, or thermostat, that constantly checks the body's core temperature to ensure that it is within a specific range. Small increases in body temperature above a certain upper threshold trigger the brain to induce disseminated peripheral cutaneous vasodilation and sweating to dissipate heat and to lower the body's temperature again. In many menopausal women, for reasons related to beta-estradiol fluctuation not yet completely understood, this upper threshold of the brain's thermostat is lowered thus leading the brain to trigger heat dissipating mechanisms several times during the day and night, causing the repeated episodes of hot flushes that are typically reported by women in this period of their lives. The relationship between the luteinising hormone (LH) surge and the brain's lowering of the sweating threshold is not yet completely clear (Cagnacci 2002).

Description of the intervention

The treatment of choice for moderate to severe vasomotor symptoms is estrogen therapy with or without a progestogen (NAMS 2012; Speroff 2011). For healthy women with annoying vasomotor symptoms, especially for those under 60 years of age and within the first 10 years after menopause, hormone therapy (HT) is still a reasonable choice. Physicians are advised to use the smallest effective dose for the shortest duration possible (Shifren 2010). Various preparations of HT can be prescribed for women in menopausal transition, including bioidentical hormone therapy (BHT). According to the Endocrine Society, a biodentical hormone is a compound that is "identical in structure to that which is produced in the human body" (ES 2006).

Many well-tested, government-approved, brand-name HT products are bioidentical hormones. Bioidentical beta-estradiol is aproved by the US Food and Drug Administration (FDA), is derived from plant sources and is available in the form of pills, patches, nasal sprays, creams, gels, and vaginal tablets. Bioidentical progesterone is also available in FDA-approved preparations (such as oral micronised progesterone in oil or vaginal progesterone gel) (Shifren 2007; Sturdee 2011). The European Medicines Agency Pre-authorisation Evaluation of Medicines for Human Use (EMEA 2005) recommends beta-estradiol alone or a combination of beta-estradiol plus progestogen for the treatment of beta-estradiol deficiency symptoms in postmenopausal women.

The findings of the combined HT arm of the Women's Health Initiative (WHI) trial raised serious concerns about the safety of HT. Combined HT (conjugated equine estrogen (CEE) 0.625 mg/d plus medroxyprogesterone acetate (MPA) 2.5 mg/d) was compared with placebo, and women in the intervention arm had increased rates of breast cancer, coronary heart disease (CHD), stroke and pulmonary embolus (Rossouw 2002). The WHI trial showed that HT is not suitable for long-term prevention of CHD and that for

any type of HT therapy the baseline risk profile of the individual woman must be taken into account (Rossouw 2013). These findings changed attitudes to HT and many women stopped taking it. Studies have shown that perimenopausal and menopausal women perceive complementary therapies as being safer and more effective because they are more 'natural' (Adams 2001; Kaufert 1998; Seidl 1998). Some opted for alternative therapies such as foods or supplements enriched with phytoestrogens (Hersh 2003, Lethaby 2013). Similarly, many women regard BHT as a safer alternative to other forms of HT (Drisko 2000; NAMS 2012).

How the intervention might work

Vasomotor symptoms associated with the onset of menopause have been shown to decline in a linear fashion as estrogen levels are elevated with replacement therapy (Corson 1993), and in this respect BHT has the same action as other forms of HT.

Why it is important to do this review

Diminished sleep quality, irritability, difficulty concentrating, and subsequently reduced quality of life (QOL) are potential consequences of the vasomotor symptoms typical of the menopausal transition (NAMS 2012). Some common HTs include synthetic oestrogens alone or combined with progestogens, and these options have benefits and risks. The use of BHT has escalated in recent years. It is necessary to look at the best available evidence on the effectiveness and safety of BHT. This review aims to assess the effectiveness and safety of the specific subgroup of bioidentical beta-estradiol and progesterone formulations for the relief of vasomotor symptoms (NAMS 2012).

OBJECTIVES

To determine the effectiveness and safety of bioidentical hormones compared to placebo or non-bioidentical hormones for the relief of vasomotor symptoms.

METHODS

Criteria for considering studies for this review

Types of studies

Randomised controlled trials (RCTs) comparing BHT for the relief of vasomotor symptoms with placebo or with non-bioidentical hormones.

Types of participants

Women in the menopausal transition (spontaneous or surgical menopause) with vasomotor symptoms.

Types of interventions

Prescription of bioidentical compared with prescription of non-bioidentical hormones.

Bioidentical hormones are compounds with a molecular and chemical structure identical to hormones produced by the ovary. We included in this review bioidentical estrogen (beta-estradiol) alone or in combination with progesterone, used in any dose or route of administration in the first group.

In the comparison group we included women receiving non-bioidentical hormones, such as equine products, estradiol valerate, estropipate (piperazine estrone sulfate) or ethinyl estradiol alone or in combination with various progestogens (medroxyprogesterone acetate, norethindrone, norethindrone acetate, drospirenone, dienogest) administered in cyclic or continuous regimens. Other drugs used for the relief of vasomotor symptoms were not included in this review.

Types of outcome measures

Primary outcomes

1. Frequency or intensity of vasomotor menopausal symptoms (hot flushes or night sweats) measured by any validated scale.

Secondary outcomes

2. Incidence and severity of adverse effects.

3. Quality of life evaluated with any validated instruments used for quality of life measures, such as the Menopause-Specific Quality of Life (Hilditch 1996), Women's Health Questionnaire (Hunter 1992), or other generic measures.

Search methods for identification of studies

Electronic searches

We searched the following sources: Cochrane Central Register of Controlled Trials (CENTRAL; 2015 Issue 7), PubMed (from inception to 29th July 2015), Embase (from inception to 29th July 2015), Literatura Latino-Americana e do Caribe em Ciências da Saúde (LILACS) (from 1982 to 29th July 2015) in consultation with the Brazilian Cochrane Centre, see Appendix 1, Appendix 2, Appendix 3, Appendix 4.

We searched for ongoing trials in the following websites:

- metaRegister of Controlled Trials (www.controlled-trials.com);

- US National Institutes of Health ongoing trials register (www.clinicaltrials.gov);
- World Health Organization International Clinical Trials Registry Platform (www.who.int/trialsearch);
- EU Clinical Trials Register (https://www.clinicaltrialsregister.eu/).

There were no language restrictions and we planned to seek translations if necessary.

Searching other resources

We screened the reference lists of the included studies for additional potentially relevant studies and contacted specialists in the field for any possible unpublished data.

Data collection and analysis

Selection of studies

Two review authors (AMISG, SSS) independently screened the trials identified by the literature search. After merging the search results and removing duplicate records, the review authors examined titles and abstracts to select the relevant reports. They then retrieved and examined the full texts of selected studies for compliance with eligibility criteria and documented the reasons for exclusion of individual trials. They consulted a third author (EMKS or CRM) if any disagreements arose (at this or at any other stage as listed below).

Data extraction and management

Two review authors (AMISG, SSS) extracted data independently using an extraction form designed and pilot-tested by the review author team. Where studies had multiple publications, we used the main trial report as the reference and additional details were supplemented from secondary papers. The review authors resolved disagreements by consensus or by discussion with a third author. We contacted the authors from all studies with incomplete information on outcomes of interest or additional missing details.

Assessment of risk of bias in included studies

We assessed the included studies for risk of bias using the 'criteria for judging risk of bias' outlined in the Cochrane tool for assessing risk of bias (Higgins 2011a). The analysis included the following: sequence generation; allocation concealment; blinding of participants, providers and outcome assessors; incomplete outcome data; selective outcome reporting; and other potential sources of bias. Two review authors independently assessed risk of bias. They resolved disagreements by consensus or discussion with a third author. All judgments were fully described and justified.

Measures of treatment effect

We expressed dichotomous data as odds ratios (OR) with 95% confidence intervals (CIs). For continuous data, we calculated mean differences (MD) and 95% CIs between treatment groups if studies reported exactly the same outcomes. If similar outcomes were reported on different scales, we calculated the standardised mean difference (SMD) and 95% CI.

Data were collected at three months or at the end of the study, or both, for the following outcomes:

- hot flush frequency - continuous;
- hot flush severity
    - continuous,
    - dichotomous (present versus absent, moderate to severe versus mild to absent) or
    - categorical (i.e. the number of women in each severity category);
- quality of life scores - continuous; and
- adverse events - dichotomous.

Where data to calculate SMDs or ORs were not available, we used the most detailed numerical data available that might facilitate a similar analysis of included studies (for example test statistics, P values).

When studies reported sufficient detail to calculate mean difference but gave no information on the associated standard deviation (SD), we assumed the outcome to have a SD equal to the highest SD from other studies using the same assessment scale. We compared the magnitude and direction of effects reported by studies with how they were presented in the review, taking into account legitimate differences. We included both data reported as final mean scores in each group and mean change scores from baseline in each group.

Where studies reported both values, we preferentially included mean change scores from baseline. For the purpose of interpretation, we considered a SMD between 0.2 and 0.5 as a small effect; between 0.5 and 0.8 as a moderate effect; and higher than 0.8 as a large effect. We reported data unsuitable for analysis in additional tables.

Unit of analysis issues

The unit of analysis was the individual participant (unit to be randomised for interventions to be compared), that is, the number of observations in the analysis matched the number of individuals randomised.

Dealing with missing data

In cases of missing outcomes, or any uncertainty regarding the data, we contacted the authors asking for the missing data or clarification. In the case of no response, irrespective of the type of data, we reported dropout rates in the 'Characteristics of included studies' table and used intention-to-treat analyses. We imputed only

the missing data of the primary outcome (vasomotor symptoms) with replacement values. For dichotomous outcomes we assumed the missing data to be treatment failures, and for continuous outcomes we imputed the mean observed. In the absence of standard deviations, we calculated it when possible from standard errors, confidence intervals, and P values for differences in means according the Cochrane Handbook (Higgins 2011b).
We performed sensitivity analyses by excluding the participants with missing data to assess the strength of the results.

Assessment of heterogeneity

We quantified inconsistency among the pooled estimates using the I^2 = ((Q - df)/Q) x 100%) statistic, where Q is the Chi^2 statistic and df represents the degree of freedom. This illustrates the percentage of the variability in effect estimates resulting from heterogeneity rather than sampling error (Deeks 2011).
The thresholds for the interpretation of I^2 were as follows:

- 0% to 25%, low heterogeneity;
- 25% to 75%, moderate;
- more than 75%, substantial heterogeneity (Higgins 2003).

Assessment of reporting biases

We planned to minimise the potential impact of publication bias and other reporting biases by ensuring a comprehensive search for eligible studies, and by being alert for duplication of data. If 10 or more studies were included in an analysis, we planned to use a funnel plot to investigate the possibility of small study effects (a tendency for the intervention to have a bigger impact in smaller studies).

Data synthesis

If no substantial heterogeneity was identified, we performed pooled estimates of the treatment effect for each outcome under a fixed-effect model. If substantial heterogeneity was identified, we performed a random-effects model analysis. The analyses were stratified by dose of BHT. When there were several intervention groups with different doses and a single control group, we divided the control group data to allow pooling without double-counting of data, and gave details in a footnote to the analysis. (Analysis 5.1; Analysis 5.2)
We conducted separate comparisons according the route of administration of the intervention.

Subgroup analysis and investigation of heterogeneity

If substantial heterogeneity was found, and there were sufficient data, we planned to explore the possible causes by using subgroup analyses. If data were available, we intended to conduct the following subgroup analyses.

- Type of menopause: natural; induced; not specified.
- Participants' baseline status: women with vasomotor symptoms; women with other associated symptoms.

As noted above, we conducted separate comparisons according the route of administration of the intervention.

Sensitivity analysis

We planned to conduct sensitivity analyses for the primary outcomes to determine whether the conclusions were robust to arbitrary decisions made during the review process regarding the eligibility for end analysis.
These analyses included consideration of whether the review conclusions would have differed if eligibility was restricted to studies without high risk of bias (studies with high risk of bias for one or more key domains).

Overall quality of the body of evidence: 'Summary of findings' table

We created a 'Summary of findings' table using GRADEpro Guideline Development Tool (GRADEpro GDT 2015). This table evaluated the overall quality of the body of evidence for the main review outcomes (vasomotor menopausal symptoms and adverse effects) and the most clinically relevant comparisons, using GRADE criteria (study limitations (i.e. risk of bias), consistency of effect, imprecision, indirectness and publication bias). Judgements about evidence quality (high, moderate or low) was justified, documented, and incorporated into reporting of results for each outcome.

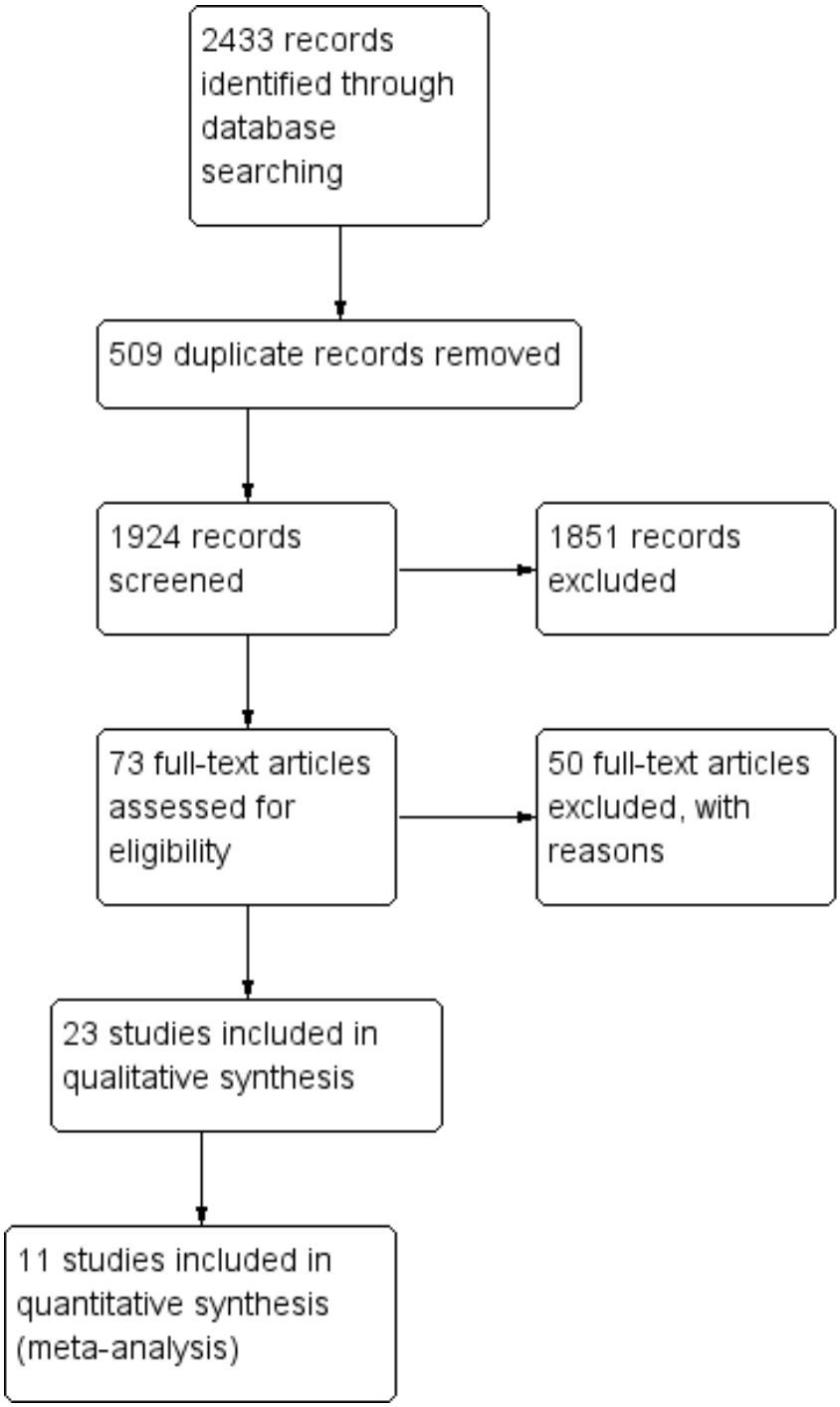
RESULTS

Description of studies

Results of the search

The last search was performed on 29 July 2015 and yielded 2433 records which were reduced to 1923 after the exclusion of duplicates. We excluded a total of 1851 references based on the title and abstract and selected 72 for full-text reading: we excluded 50 of these because they did not fulfil the selection criteria, and 23 studies were included in the review (Figure 1).

Figure 1.



Included studies

Trial design

All of the trials were double-blind, placebo-controlled, randomised clinical trials. Most of the included studies were multicentre studies conducted in the USA (Archer 1992; Archer 2003; Archer 2012; Buster 2008; Cohen 1999; Good 1999; Gordon 1995; Notelovitz 2000a; Notelovitz 2000b; Simon 2006; Speroff 1996a; Speroff 1996b; Utian 1999) or USA and Canada (Simon 2007). Three studies were conducted in Italy (Bacchi-Modena 1997; De Aloysio 2000; Rovati 2000) and the others were conducted in Asia (Haines 2009), Denmark (Nielsen 2006), Germany (von Holst 2000), Japan (Honjo 2009), Sweden (Wiklund 1993), and The Netherlands (De Vrijer 2000).

Four studies contained statements declaring that they were supported by the pharmaceutical industry (Good 1999; Notelovitz 2000a; Speroff 1996a; Speroff 1996b)

Participants

The studies included a total of 5779 women, who were either naturally or surgically (bilateral oophorectomy) menopausal. Most of the studies (20/23) included only women with moderate to severe vasomotor symptoms.

The duration of natural menopause ranged from at least six months (Archer 2003; Archer 2012;Buster 2008; De Aloysio 2000; Good 1999; Notelovitz 2000a; Rovati 2000; Wiklund 1993) to eight to 12 months (Bacchi-Modena 1997; Cohen 1999; De Vrijer 2000; Gordon 1995; Honjo 2009; Haines 2009; Notelovitz 2000b; Simon 2006; Simon 2007; Utian 1999) and up to five years (Nielsen 2006). Three studies (Speroff 1996a; Speroff 1996b; Archer 1992) did not report the time undergone since natural menopause of their participants. For the surgically menopausal participants, the time since bilateral oophorectomy ranged from two to four weeks (De Aloysio 2000; Gordon 1995; Honjo 2009; Notelovitz 2000b; Speroff 1996a) to at least six weeks (Bacchi-Modena 1997; Cohen 1999; De Vrijer 2000; Nielsen 2006; Haines 2009; Rovati 2000; Simon 2007; Utian 1999 ; von Holst 2000; Wiklund 1993) up to one year or more before study entry (Simon 2006). Six studies, (Archer 1992; Archer 2003; Archer 2012; Buster 2008; Good 1999; Notelovitz 2000a) did not report the time between surgical menopause up to study entry.

Archer 2003, Good 1999, De Vrijer 2000, De Aloysio 2000, Notelovitz 2000b and Utian 1999 excluded women who had received oral, transdermal, or vaginal steroid hormones (beta-estradiols, progestogens or corticosteroids) in the last six to eight weeks prior to enrolment. De Vrijer 2000, De Aloysio 2000, Gordon 1995 and Rovati 2000 stated that they excluded women who had received implants in the 12 preceding months, or with any previous unopposed beta-estradiol use for more than three months. Bacchi-Modena 1997, Cohen 1999, Simon 2006 and von Holst 2000 excluded women who had used transdermal HT in the last 21 days to two months or oral hormones in the last 28 days to eight weeks, or six weeks vaginal beta-estradiol, or injection or implants in the three to six months preceding trial entry. The other studies (Archer 1992; Archer 2012; Buster 2008; Honjo 2009; Haines 2009; Nielsen 2006; Notelovitz 2000a; Simon 2007; Speroff 1996a; Speroff 1996b; Utian 1999; Wiklund 1993) did not specify washout periods for the participants before entry in the trials.

All the trials reported confirmation of ovarian failure by measurement of FSH levels. Postmenopausal status was confirmed through serum estradiol concentration < 20 pg/mL, serum FSH level at least 40-50 mIU/mL and a minimum of 3-7 hot flushes per day or 56-60 hot flushes per week of moderate to severe intensity. Nielsen 2006 included women who were fewer than five years past menopause at study entry and had serum estradiol levels below 0.16 nmol/L and FSH levels above 42 IU/L.

We have presented full details of the inclusion and exclusion criteria in the 'Characteristics of included studies' table.

Interventions

Eleven studies compared the effects of transdermal beta-estradiol patches (0.014, 0.02, 0.025, 0.0375, 0.04, 0.05, 0.1 and 0.2 mg/d) versus placebo on hot flush frequency, severity, or both (Bacchi-Modena 1997; Cohen 1999; De Aloysio 2000, De Vrijer 2000; Haines 2009; Rovati 2000; Speroff 1996a; Speroff 1996b; Utian 1999; von Holst 2000; Wiklund 1993).

Three studies (Archer 2003; Archer 2012; Simon 2007) used beta-estradiol gel (0.27 mg/d; 0.37 mg/d; 0.5 mg/d; 0.75 mg/d; 1.0 mg/d and 1.5 mg/d).“

Four studies compared oral beta-estradiol (0.1 mg/d; 0.25 mg/d; 0.5 mg/d; 1.0 mg/d; 1.2 mg/d and 2 mg/d) with placebo (Archer 1992; Honjo 2009; Notelovitz 2000a; Notelovitz 2000b).

One study compared micellar nanoparticle beta-estradiol emulsion versus placebo (Simon 2006).

Two studies (Buster 2008; Nielsen 2006) used intranasal spray (0.021 mg/d; 0.029 mg/d; 0.040 mg/d; 150 µg/d; and 300 µg/d) as the mode of beta-estradiol delivery.

Three studies compared transdermal beta-estradiol (0.05 or 0.1 mg/d) or oral beta-estradiol (1 or 2 mg/d) versus 0.625 or 1.25 mg/d oral conjugated equine estradiol (CEE) (Archer 1992; Good 1999; Gordon 1995).

Outcomes

The trials measured the primary outcome (vasomotor symptoms) in many different ways. None of the included studies reported night sweats as a separate outcome. Most commonly, study participants recorded the number of episodes over a period of a day or week, and changes from the baseline indicated treatment response.
Seventeen studies assessed vasomotor symptoms such as hot flush frequency and severity (Archer 2003; Archer 2012; Buster 2008; Bacchi-Modena 1997; Cohen 1999; De Aloysio 2000; De Vrijer 2000; Good 1999; Gordon 1995; Honjo 2009; Notelovitz 2000a; Notelovitz 2000b; Rovati 2000; Simon 2006; Simon 2007; Utian 1999; von Holst 2000).
Four studies assessed only hot flush frequency (Archer 1992; Haines 2009; Speroff 1996a; Speroff 1996b).
Climacteric symptoms were assessed by the Greene Climacteric Scale (Buster 2008); Kupperman's Index (Bacchi-Modena 1997; De Aloysio 2000; Rovati 2000; von Holst 2000; Wiklund 1993); and visual analogue scale (VAS). One study did not report data using a specific vasomotor symptom subscale (Honjo 2009).
Severity was evaluated according to different rating scales ranging from 1 to 3 (mild to severe symptoms) (Archer 2003; Archer 2012; Buster 2008; Cohen 1999; De Vrijer 2000; Good 1999; Gordon 1995; Honjo 2009; Simon 2006; Simon 2007; Speroff 1996a; Speroff 1996b; Notelovitz 2000a; Notelovitz 2000b; Utian 1999; von Holst 2000) by a VAS scale of 0 to 100 (Bacchi-Modena 1997; De Aloysio 2000; Rovati 2000) and according to log transformation of the severity scores obtained.
Most studies reported the mean reduction in average 24-hour hot flush frequency and severity from baseline to the end or any time of the study (Archer 1992; Archer 2003; Archer 2012; Bacchi-Modena 1997; Buster 2008; Cohen 1999; De Aloysio 2000; De Vrijer 2000; Honjo 2009; Notelovitz 2000b; Rovati 2000; Simon 2006; Simon 2007; Utian 1999). Several studies assessed the mean weekly reduction in frequency and severity of moderate to severe hot flushes from baseline (Good 1999; Gordon 1995; Haines 2009; Notelovitz 2000a; von Holst 2000). Two studies used analysis of covariance to compare the frequency of hot flushes per week (Speroff 1996a; Speroff 1996b). The vasomotor symptoms diary was the tool that the included studies used most frequently to quantify hot flush frequency and severity. For all RCTs, the frequency of hot flushes was based on the number of hot flushes recorded per day.
Only three studies reported quality of life as one of their outcomes (Haines 2009; Nielsen 2006; Wiklund 1993).
All studies reported adverse events.
We imputed SDs for two studies that failed to report them (Bacchi-Modena 1997; von Holst 2000).
Fourteen of the 23 studies reported data that were unsuitable for analysis for our primary outcome because data were skewed or were presented only in graphical form (Archer 1992; Archer 2003; Archer 2012; De Aloysio 2000; Good 1999; Gordon 1995; Haines 2009; Nielsen 2006; Notelovitz 2000a; Rovati 2000; Speroff 1996a; Speroff 1996b; Simon 2007; Simon 2006). We have reported their findings in additional tables.

Excluded studies

Fifty studies were excluded mainly because they did not include the intervention or any of the outcomes of interest, or because of their design. See Characteristics of excluded studies for more details.

Risk of bias in included studies

See Figure 2 and Figure 3 for the graph and summary of review authors' judgements about the risk of bias for included studies.

Figure 2.

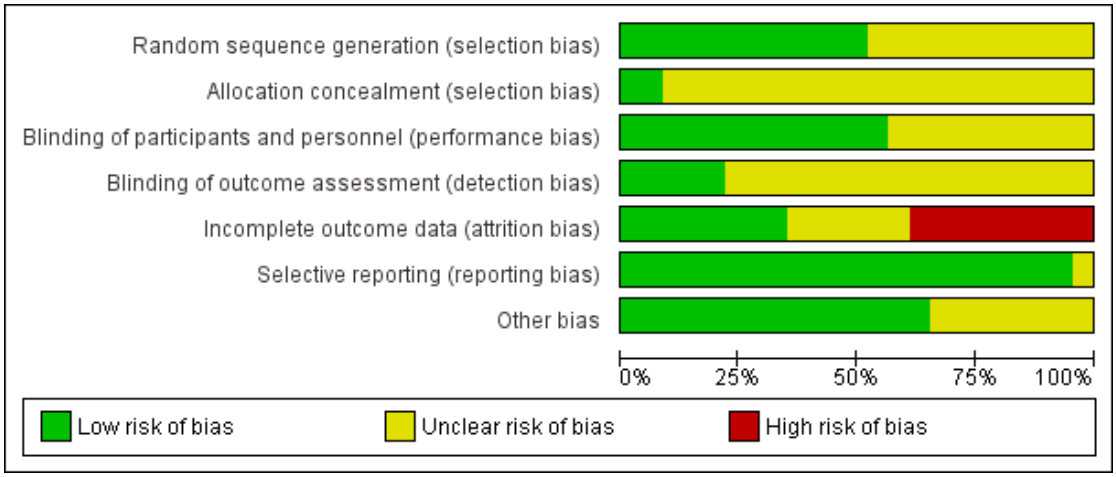
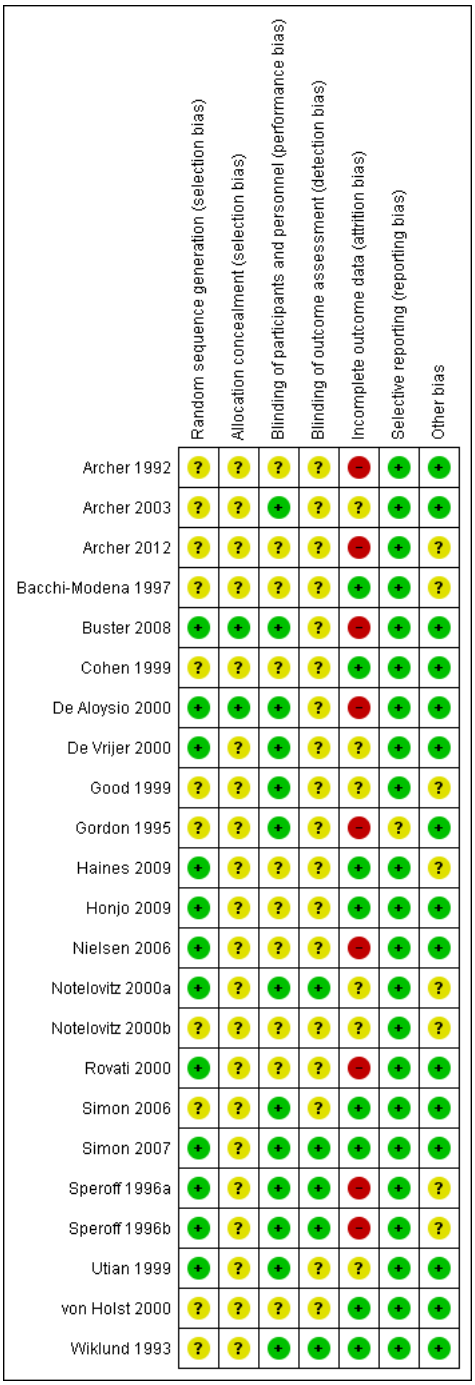


Figure 3. Risk of bias summary: review authors' judgements about each risk of bias item for each included study.



Allocation

Sequence generation

Eleven studies (Archer 1992; Archer 2003; Archer 2012; Bacchi-Modena 1997; Cohen 1999; Good 1999; Gordon 1995; Notelovitz 2000b Simon 2006; von Holst 2000; ; Wiklund 1993) were at unclear risk of selection bias because they did not describe the method of sequence generation. The other studies (Buster 2008; De Aloysio 2000; De Vrijer 2000; Haines 2009; Honjo 2009; Nielsen 2006; Notelovitz 2000a; Rovati 2000; Simon 2007; Speroff 1996a; Speroff 1996b, Utian 1999) reported adequate methods of sequence generation and we deemed them to be at low risk of bias.

Allocation concealment

Eighteen studies (Archer 1992; Archer 2003; Archer 2012; Bacchi-Modena 1997; Cohen 1999; De Vrijer 2000; Good 1999; Honjo 2009; Haines 2009; Notelovitz 2000a; Notelovitz 2000b; Rovati 2000; Simon 2006; Speroff 1996a; Speroff 1996b; Utian 1999; von Holst 2000; Wiklund 1993) did not report the method used for allocation concealment and thus we considered them at unclear risk of selection bias. Buster 2008; De Aloysio 2000; Gordon 1995; Nielsen 2006; Simon 2007 reported adequate allocation concealment methods and we therefore classified them as being at low risk of bias for this domain.

Blinding

Performance bias

All 23 studies reported that the participants were blinded but only 13 studies provided sufficient details about the blinding (Archer 2003; Buster 2008; De Aloysio 2000; De Vrijer 2000; Good 1999; Gordon 1995; Notelovitz 2000a; Simon 2006; Simon 2007; Speroff 1996a; Speroff 1996b; Utian 1999; Wiklund 1993) and we rated them at low risk of performance bias. The other studies we rated at unclear risk of bias in this domain.

Detection bias

Five studies provided adequate detail about blinding of outcome assessment and we rated them at low risk of detection bias (Notelovitz 2000a; Simon 2007; Speroff 1996a; Speroff 1996b; Wiklund 1993). The other studies described themselves as double-blinded but reported inadequate information about blinding of outcome assessment, and we rated them at unclear risk of detection bias.

Incomplete outcome data

We rated studies at high risk of attrition bias if 20% or more participants were lost to follow-up, or if losses were over 10% and unbalanced between the groups. We rated studies at unclear risk of attrition bias if 10% to 20% of participants were not included in analysis, and at low risk if fewer than 10% were not included in analysis.

All studies had loss of participants at follow-up except for one study which did not provide information on this (De Vrijer 2000). We deemed eight studies at low risk of bias for this domain (Bacchi-Modena 1997; Cohen 1999; Haines 2009; Honjo 2009; Simon 2006; Simon 2007; von Holst 2000; Wiklund 1993). Three studies reported that losses to follow-up were low and balanced in all groups (Bacchi-Modena 1997; Cohen 1999; Utian 1999). We considered nine studies to be at high risk of attrition bias because the dropout rates were high or not balanced between the groups, or both (Archer 1992; Archer 2012; Buster 2008; De Aloysio 2000; Gordon 1995; Nielsen 2006; Rovati 2000; Speroff 1996a; Speroff 1996b). The other studies were deemed at unclear risk.

Selective reporting

We rated studies at low risk of selective reporting if they reported all expected outcomes (vasomotor symptoms and adverse events), including those prespecified in the protocol or methods sections. Twenty-two studies included all expected outcomes and we deemed them to be at low risk of reporting bias. We rated risk of bias for Gordon 1995 as unclear because reporting of adverse events was limited.

Other potential sources of bias

We rated eight studies at unclear risk of bias due to failure to report full statistical data, either because data were only presented as a graph or because no standard deviations were reported. All other studies were rated at low risk of bias in this domain, as we did not identify any other source of potential bias.

Effects of interventions

See: **Summary of findings for the main comparison** Transdermal beta-estradiol patch versus placebo for women with hot flushes; **Summary of findings 2** Transdermal beta-estradiol gel versus

placebo for women with hot flushes; **Summary of findings 3** Oral beta-estradiol versus placebo for women with hot flushes; **Summary of findings 4** Topical beta-estradiol emulsion versus placebo for women with hot flushes; **Summary of findings 5** Intranasal beta-estradiol versus placebo for women with hot flushes; **Summary of findings 6** Transdermal beta-estradiol patch versus conjugated equine estrogens for women with hot flushes; **Summary of findings 7** Oral beta-estradiol versus conjugated equine estrogens for women with hot flushes

The included studies evaluated the effects of BHT given through different routes of administration and in different dosages, making it difficult to conduct meta-analyses. When it was not possible to perform meta-analyses, we presented a narrative description of the results.

We present the comparisons as follows:

- **Beta-estradiol versus placebo**
  - Transdermal beta-estradiol patch versus placebo (11 RCTs)
  - Transdermal beta-estradiol gel versus placebo (three RCTs)
  - Oral beta-estradiol versus placebo (four RCTs)
  - Topical beta-estradiol emulsion versus placebo (one RCT)
  - Intranasal beta-estradiol versus placebo (two RCTs)
- **Beta-estradiol versus conjugated equine** estrogen**s (CEE)**
  - Transdermal beta-estradiol patch versus CEE (two RCTs)
  - Oral beta-estradiol versus CEE (one RCT)

1. Beta-estradiol versus placebo

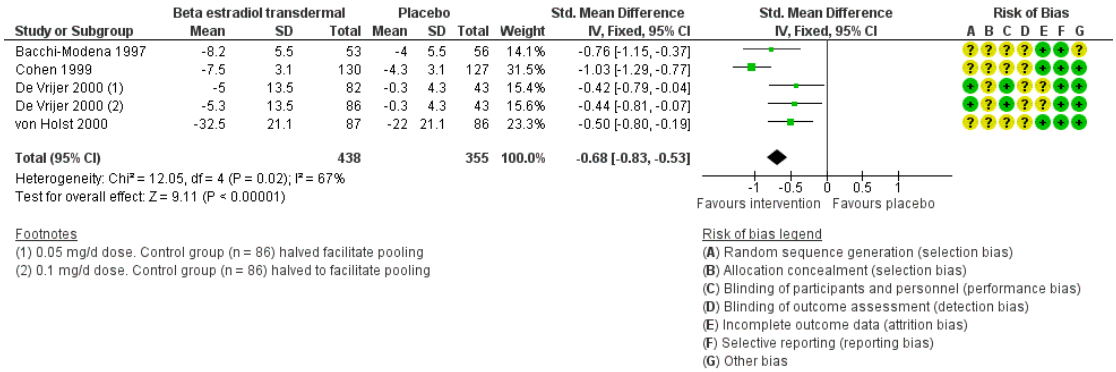
1.1 Transdermal beta-estradiol patch versus placebo (11 RCTs)

1.1.1 Primary outcome: frequency or intensity of vasomotor menopausal symptoms measured by any validated scale

Frequency of hot flushes

Four RCTs reported data suitable for pooling. The daily dose of beta-estradiol was 0.0375 mg/d (Cohen 1999), 0.05 mg/d (Bacchi-Modena 1997; De Vrijer 2000; von Holst 2000) or 0.10 mg/d (De Vrijer 2000). The SMD analysis was applied due to differences in the measurement of hot flush frequency (reduction in the number of hot flushes per week, per day). Our findings were compatible with a moderate to large benefit in the intervention group (SMD -0.68, 95% CI -0.83 to -0.53, four RCTs, 793 women, I^2 = 67%, low quality evidence). There was moderate statistical heterogeneity which was attributable to one study (Cohen 1999), in which the benefit in the intervention group was more pronounced; though there was no obvious difference between Cohen 1999 and the other studies that might account for this. However, the direction of effect was consistent (Analysis 1.1, Figure 4).

Figure 4. Forest plot of comparison: 1 Transdermal beta estradiol patch vs placebo, outcome: 1.1 Frequency of hot flushes



Subgroup analysis by dose

All doses of beta-estradiol were associated with a benefit for the intervention group. There was evidence of a significant difference between the subgroups (test for subgroup differences: Chi^2 = 10.33, df = 2 (P value = 0.006), I^2 = 80.6%). The difference

was attributable to Cohen 1999. It was unclear why the effect was more pronounced in this study, as it utilised the lowest dose of beta-estradiol.

Findings in studies with data unsuitable for analysis

We were unable to extract data suitable for analysis in seven studies. They utilised a daily dose of beta-estradiol of 0.014 mg (Haines 2009), 0.02 mg (Speroff 1996a; Speroff 1996b) 0.025 mg (De Aloysio 2000; Rovati 2000; Utian 1999), 0.375 mg (De Aloysio 2000), 0.04 mg (Speroff 1996a, Speroff 1996b,) 0.05 mg (Gordon 1995; Rovati 2000; Utian 1999) or 0.1 mg (Gordon 1995; Utian 1999). In all comparisons there was a benefit in the intervention group (see Table 1 for details).

Intensity of hot flushes

Two RCTs reported data suitable for pooling. The daily dose of beta-estradiol was 0.025 mg (De Aloysio 2000; Rovati 2000), 0.0375 mg (De Aloysio 2000) or 0.05 mg (Rovati 2000). Measured on a 0 to 100 visual analogue scale (VAS), the intensity of hot flushes was lower in the intervention group (MD -19.94 points, 95% CI -24.86 to -15.02, two RCTs, 393 women, I^2 = 54%, low quality evidence). There was moderate statistical heterogeneity, but the direction of effect was consistent (Analysis 1.2).

Subgroup analysis by dose

All doses of beta-estradiol were associated with a benefit for the intervention group. There was evidence of a significant difference between the subgroups (test for subgroup differences: Chi^2 = 6.29, df = 2, P value = 0.04, I^2 = 68.2%), which was attributable to the subgroup which received the highest dose (0.05 mg/d), among whom the benefit in the intervention group was more pronounced.

Findings in studies with data unsuitable for analysis

We were unable to extract data suitable for analysis in one study, which compared a weekly 0.014 mg/d transdermal 17 beta-estradiol patch versus placebo (Haines 2009). The relative reduction was larger in the intervention than in the placebo group: -5.8 versus -8.4, P value < 0.05 (see Table 2 for details).

Secondary outcomes

1.1.2 Incidence and severity of adverse effects

Nine RCTs reported data suitable for pooling. The daily dose of beta-estradiol was 0.014 mg (Haines 2009), 0.025 mg (De Aloysio 2000; Rovati 2000; Utian 1999), 0.0375 mg (Cohen 1999; De Aloysio 2000), 0.05 mg (Bacchi-Modena 1997; De Vrijer 2000; Rovati 2000; Utian 1999; von Holst 2000; Wiklund 1993) or 0.10 mg (De Vrijer 2000; Utian 1999).

Adverse events (such as skin irritation, vaginal bleeding and breast tenderness) were more common in the intervention arm (OR 2.14, 95% CI 1.29 to 3.54, 9 RCTs, 1822 women, I^2 = 73%, low quality evidence). There was moderate statistical heterogeneity, but the direction of effect was consistent (Analysis 1.3). In one study (De Vrijer 2000), five women in the intervention group developed endometrial hyperplasia. The heterogeneity was mainly attributable to the subgroup that had the 0.10 dose of beta-estradiol.

Subgroup analysis by dose

When subgrouped by dose, findings were no longer statistically significant except for the subgroup of women who had the highest dose of beta-estradiol (0.10 mg). There was evidence of a significant difference between the subgroups (test for subgroup differences: Chi^2 = 27.48, df = 4, P value < 0.00001, I^2 = 85.4%), attributable to a more pronounced effect in this high-dose subgroup.

1.1.3 Quality of life evaluated with any validated instrument

Two studies reported quality of life as an outcome, using a variety of scales. The daily dose of beta-estradiol was 0.014 mg (Haines 2009) and 0.05 mg (Wiklund 1993).

Haines 2009 reported no evidence of a difference between the groups, measured on the Menopause Quality of Life (MENQol) scale (MD 0.00 points, 95% CI -0.35 to 0.35, one RCT, 165 women) (Analysis 1.4).

Wiklund 1993 observed greater improvement in the intervention group, measured using five different scales (see Analysis 1.5).

1.2 Transdermal beta-estradiol gel versus placebo (three RCTs)

1.2.1 Primary outcome: frequency or intensity of vasomotor menopausal symptoms

There were no data suitable for analysis.

*Findings in studies with data unsuitable for analysis (*Table 1)

Three studies reported data unsuitable for analysis (Archer 2003; Archer 2012; Simon 2007). The daily dose of beta-estradiol was 0.27 mg (Archer 2012), 0.37 mg (Archer 2012), 0.5 mg (Simon 2007), 0.75 (Archer 2003), 1.0 (Simon 2007), or 1.5 mg (Archer 2003; Simon 2007). See Table 3 for details.

Archer 2003 and Archer 2012 reported a benefit in the all the intervention groups compared with the placebo groups (P value < 0.05 and P value < 0.001 respectively).

Simon 2007 also reported that the proportion of women with adequate relief of hot flushes was significantly higher in the intervention groups (P value < 0.001)

Secondary outcomes

1.2.2 Incidence and severity of adverse effects

Three RCTs reported data suitable for pooling. The daily dose of beta-estradiol was 0.27 mg (Archer 2012), 0.37 mg (Archer 2012), 0.5 mg (Simon 2007), 1.0 mg (Simon 2007), or 1.5 mg (Simon 2007).

Adverse events such as headache and breast pain were more common in the intervention arm (OR 1.41, 95% CI 1.09 to 1.83, 3 RCTs, 1086 women, I^2 = 0%, moderate quality evidence) Analysis 2.1.

Subgroup analysis by dose

When subgrouped by dose, findings were no longer statistically significant except for the subgroup of women who had the highest dose of beta-estradiol (1.5 mg). however, there was no evidence of a significant difference between the subgroups (test for subgroup differences: Chi² = 3.75, df = 5, P value = 0.59, I² = 0%).

1.2.3 Quality of life

This outcome was not reported in the included studies.

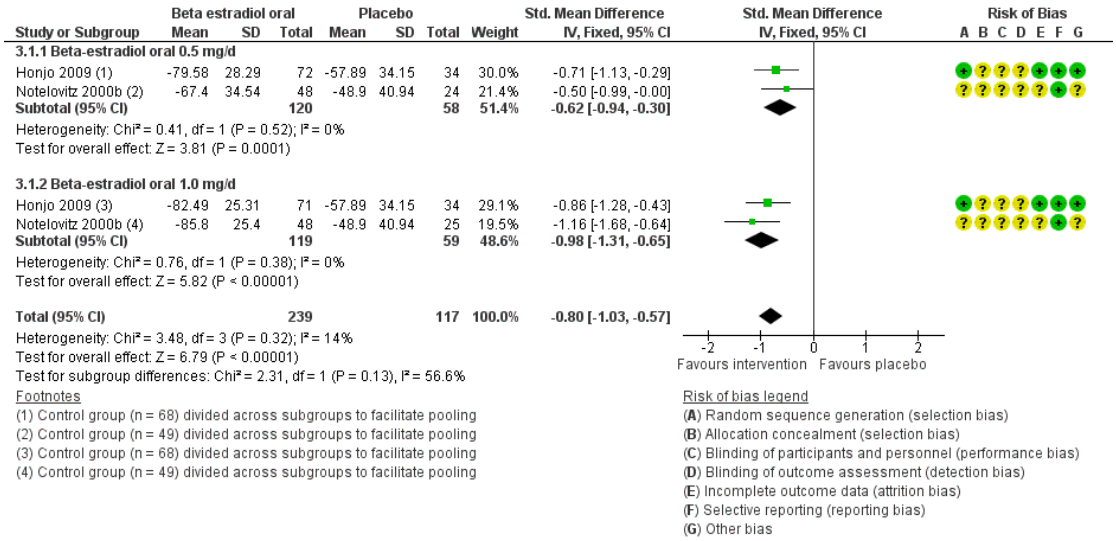
1.3 Oral beta-estradiol versus placebo (four RCTs)

1.3.1 Primary outcome: frequency or intensity of vasomotor menopausal symptoms

Hot flush frequency

Two studies reported data suitable for pooling. The daily dose of beta-estradiol was 0.5 mg (Honjo 2009; Notelovitz 2000b) or 1.0 mg (Honjo 2009; Notelovitz 2000b). The SMD analysis was applied due to differences in the scale of measures used. Our findings were compatible with a moderate to large benefit in the intervention group (SMD -0.80, 95% CI -1.03 to -0.57, 2 RCTs, 356 women, I^2 = 14%, low quality evidence) (Figure 5).

Figure 5. Forest plot of comparison: 3 Oral beta estradiol vs placebo, outcome: 3.1 Frequency of hot flushes



Subgroup analysis by dose

When data were subgrouped by dose, a benefit was seen in the intervention group at both dose levels. There was no evidence of a significant difference between the subgroups (test for subgroup differences: Chi² = 2.31, df = 1, P value = 0.13, I² = 56.6%).

Findings in studies with data unsuitable for analysis

Two studies reported data unsuitable for analysis (Archer 1992; Notelovitz 2000a).

Notelovitz 2000a compared oral micronised beta-estradiol 0.25

mg, 0.5 mg, 1 mg, or 2 mg versus placebo. Results were reported only in graphics. The authors reported a significant linear correlation between increased dosage of oral beta-estradiol and decreased moderate to severe hot flushes at week 12.

Archer 1992 compared 1.0 and 2.0 mg/d of beta-estradiol versus placebo and reported that hot flush frequency at 12 weeks was significantly lower in both intervention groups (see Table 4 for details).

Secondary outcomes

1.3.2 Incidence and severity of adverse effects

Three RCTs reported data suitable for pooling. The daily dose of beta-estradiol was 0.5 mg (Honjo 2009; Notelovitz 2000b), 1.0 mg (Archer 1992; Honjo 2009; Notelovitz 2000b) or 2 mg/d (Archer 1992).

There was no evidence of a difference between the groups (OR 1.28, 95% CI 0.84 to 1.96, 3 RCTs, 433 women, I^2 = 0%, low quality evidence) Analysis 3.2.

Subgroup analysis by dose

Subgrouping by dose did not change the statistical significance of the findings and the test for subgroup differences was not significant (test for subgroup differences: Chi² = 1.00, df = 2, P value = 0.61, I² = 0%).

1.2.3 Quality of life

This outcome was not reported in the included studies.

1.4 Topical beta-estradiol emulsion versus placebo (one RCT)

1.4.1 Primary outcome: frequency or intensity of vasomotor menopausal symptoms

One study reported this comparison (Simon 2006). The daily dose of beta-estradiol was 8.6 mg. Data were unsuitable for analysis. At the end of week 12, the intervention group had a greater reduction in the mean number of moderate to severe hot flushes per day (see Table 5 for details).

Secondary outcomes

1.4.2 Incidence and severity of adverse effects

There was no evidence of a difference between the groups (OR 1.46, 95% CI 0.80 to 2.66, 1 RCT, 200 women, low quality evidence)

1.4.3 Quality of life evaluated

This outcome was not reported

1.5 Intranasal beta-estradiol versus placebo (two RCTs)

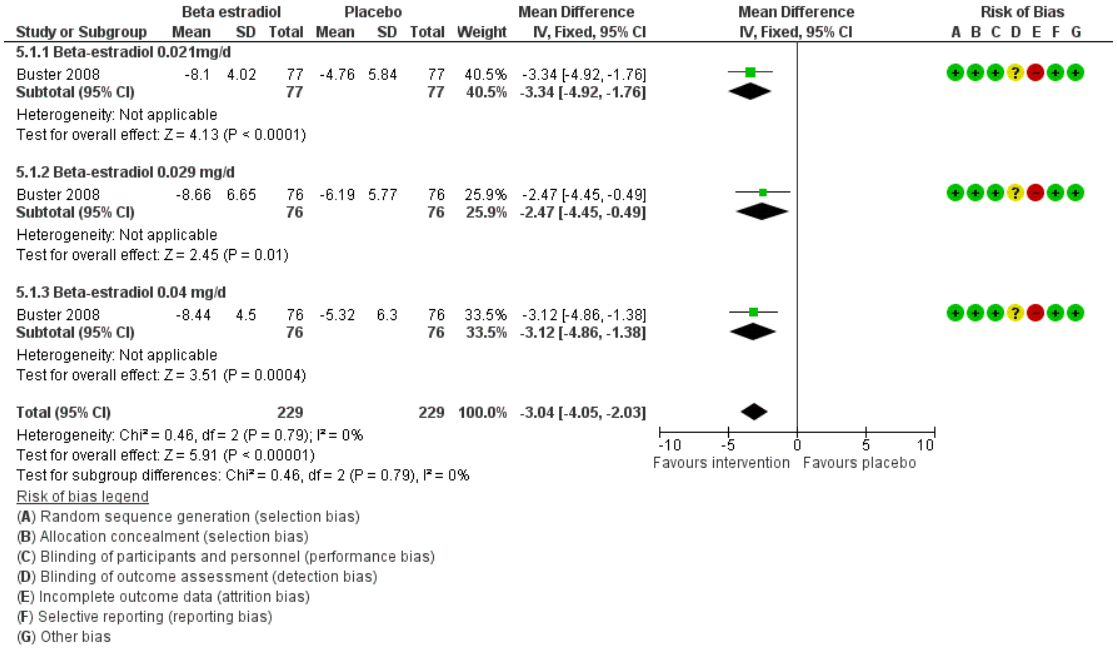
1.5.1 Primary outcome: frequency or intensity of vasomotor menopausal symptoms

Only one study reported data suitable for analysis (Buster 2008). The daily dose of beta-estradiol was one, two or three spray doses, corresponding to 0.021 mg/d, 0.029 mg/d, and 0.040 mg/d in Buster 2008. A second study reported data unsuitable for analysis (Nielsen 2006).

Hot flush frequency

There were fewer hot flushes per day in the intervention group (MD -3.04 95% CI -4.05 to -2.03, 1 RCT, 458 women, moderate quality evidence) (Analysis 5.1, Figure 6).

Figure 6. Forest plot of comparison: 5 Intranasal beta estradiol vs placebo, outcome: 5.1 Frequency of hot flushes



Subgroup analysis by dose

When data were subgrouped by dose, a benefit was seen in the intervention group at all dose levels. There was no evidence of a significant difference between the subgroups (test for subgroup differences: Chi² = 0.46, df = 2, P value = 0.79, I² = 0%).

Secondary outcomes

1.5.2 Incidence and severity of adverse effects

There was a higher rate of adverse effects (such as headache, breast tenderness, arthralgia and nausea) in the intervention group (OR 1.96, 95% CI 1.26 to 3.03, 1 RCT, 458 women).

Subgroup analysis by dose

When data were subgrouped by dose, there was no significant difference between the groups at any dose level. There was no evidence of a significant difference between the subgroups (test for subgroup differences: Chi² = 0.03, df = 2, P value = 0.99, I² = 0%).

1.5.3 Quality of life

No studies reported data suitable for analysis for this outcome

Findings in study with data unsuitable for analysis

The Nielsen 2006 study used the Women's Health Questionnaire to compare the quality of life of women using intranasal beta-estradiol 0.15 mg/d or 0.30 mg/d versus placebo.There was an improvement in both beta-estradiol groups compared to placebo in the memory and concentration, vasomotor symptoms, sleep problems and sexual behaviour dimensions (P value < 0.001). There was no evidence of a difference between the intervention and placebo groups in the anxiety, depressed mood or well-being dimensions (see Table 6 for details).

2 Beta-estradiol versus conjugated equine estrogens

2.1 Transdermal beta-estradiol patch versus CEE (two RCTs)

2.1.1 Primary outcome: frequency or intensity of vasomotor menopausal symptoms

No studies reported data suitable for analysis.

*Findings in studies with data unsuitable for analysis (*see Table 7 for details).

Two RCTs reported this outcome. They compared beta-estradiol patches with daily doses of 0.05 mg or 0.1 mg versus CEE 0.0625 mg/d (Good 1999; Gordon 1995) or 1.25 mg/d (Good 1999). Good 1999 reported that at 12 weeks there was no evidence of a

difference between any of the groups in the mean number of hot flushes per week compared with baseline (P value > 0.05).
Gordon 1995 reported that at 11 weeks there was a significant reduction in the mean number of weekly hot flushes in all three treatment groups compared with baseline: with no evidence of a difference between the groups.

Secondary outcomes

2.1.2 Incidence and severity of adverse effects

Good 1999 reported no evidence of a difference between groups in adverse events. Gordon 1995 reported that breast pain and vaginal bleeding were more frequent in the 0.1 mg/d beta-estradiol group than in other groups (P value < 0.05).

2.1.3 Quality of life evaluated with any validated instrument

This outcome was not reported.

2.2 Oral beta-estradiol versus CEE (one RCT)

2.2.1 Primary outcome: frequency or intensity of vasomotor menopausal symptoms

No studies reported data suitable for analysis
Findings in study with data unsuitable for analysis
Archer 1992 compared oral beta-estradiol 1.0 mg/d or 2.0 mg/d versus CEE 0.625 mg/d or 1.25 mg/d. After 12 weeks of use, there was no evidence of a difference between any of the groups (see Table 8 for details).

Secondary outcomes

2.2.2 Incidence and severity of adverse effects

Archer 1992 reported that there was no evidence of a difference between the groups in the incidence of possible drug-related adverse effects between the beta-estradiol 1.0 mg/d or 2 mg/d and CEE 0.625 mg/d or 1.25 mg/d (OR 1.20, 95% CI 0.50 to 2.87, 1 RCT, 103 women, I^2 = 0%). See Analysis 6.1.

2.2.3 Quality of life evaluated with any validated instrument

This outcome was not reported.

Other analyses

Sensitivity and subgroup analyses and assessment for publication bias

As there were too few studies included in any one analysis, we did not perform a sensitivity analysis to explore the robustness of the results and we did not use a funnel plot to investigate the possibility of small study effects (a tendency for the intervention to have a bigger impact in smaller studies). Nor were there sufficient data to perform subgroup analyses by type of menopause or participants' baseline status.

ADDITIONAL SUMMARY OF FINDINGS *[Explanation]*

Transdermal beta-estradiol gel versus placebo

Population: women with hot flushes
Setting: community
Intervention: beta-estradiol transdermal gel
Comparison: placebo

Outcomes	Anticipated absolute effects* (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Transdermal beta-estradiol gel versus placebo				
Frequency of hot flushes Beta-estradiol gel 0.27-1.5 mg/day	There were no data suitable for analysis. 3 studies with data unsuitable for analysis reported a benefit in the beta-estradiol gel group (P value < 0.05)		930 3 RCTs	⊕◯◯◯ **low**[1]	
Adverse effects Beta-estradiol gel dose 1.5 mg/day	Rate in placebo group: 431 per 1000 — Rate in beta-estradiol group*: 516 per 1000 (452 to 581)	OR 1.41 (1.09 to 1.83)	1086 (3 RCTs)	⊕⊕⊕◯ **moderate**[2]	The rate of adverse effects was higher in the beta-estradiol group

***The risk in the intervention group** (and its 95% confidence interval) is based on the median risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval; **OR:** Odds ratio; mg/d: milligrams per day

GRADE Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

1. Downgraded two levels for very serious risk of bias: studies at high or unclear risk of bias in most domains, data unsuitable for analysis
2. Downgraded one level for serious risk of bias: studies at high or unclear risk of bias in most domains

Oral beta-estradiol versus placebo

Population: women with hot flushes
Setting: community
Intervention: oral beta-estradiol
Comparison: placebo

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Oral beta-estradiol versus placebo for women with hot flushes					
Frequency of hot flushes Oral beta-estradiol 0.5-1.0 mg/day	There were fewer hot flushes in the beta-estradiol group. The effect size was moderate to large (SMD -0.80, 95% CI -1.03 to -0.57)			356 (2 studies)	⊕⊕◯◯ **low**[1,2]	Two studies with data unsuitable for analysis also reported a benefit in the intervention groups
Adverse effects Oral beta-estradiol 0.5-2.0 mg/day	Rate in placebo group: 245 per 1000	Rate in beta-estradiol group*: 293 per 1000 (214 to 389)	OR 1.28 (0.84 to 1.96)	433 (3 RCTs)	⊕⊕◯◯ **low**[1,2]	There was no evidence of a difference between the groups

***The risk in the intervention group** (and its 95% confidence interval) is based on the median risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval; SMD: Standardised mean difference **OR:** Odds ratio; mg/d: milligrams per day

GRADE Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

1. Downgraded one level for serious risk of bias: studies at high or unclear risk of bias in most domains
2. Downgraded one level for serious imprecision: small overall sample size and/or wide confidence interval

Topical micellar nanoparticle beta-estradiol emulsion versus placebo

Population: women with vasomotor symptoms
Setting: community
Intervention: topical micellar nanoparticle beta-estradiol emulsion
Comparison: placebo

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with Placebo	Risk with Topical micellar nanoparticle beta-estradiol				
Frequency of vasomotor menopausal symptoms Micellar nanoparticle beta-estradiol 8.6 mg/day	One study with data unsuitable for analysis reported a benefit in the beta-estradiol group (P value < 0.001)			200 (1 study)	⊕⊕◯◯ **low**[1,2]	
Adverse effects Micellar nanoparticle beta-estradiol 8.6 mg/day	650 per 1000	731 per 1000 (598 to 832)	OR 1.46 (0.80 to 2.66)	200 (1 study)	⊕⊕◯◯ **low**[1,3]	

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the median risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

CI: Confidence interval; **OR:** Odds ratio;

GRADE Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

[1]Downgraded two levels for very serious risk of bias: methods of randomisation not clearly reported, data skewed and unsuitable for analysis

[2]Downgraded one level for serious risk of bias: methods of randomisation not clearly reported

[3]Downgraded one level for serious imprecision: findings compatible with benefit in either group or with no difference between the groups

Intranasal beta-estradiol versus placebo

Patient or population: women with hot flushes
Setting: community
Intervention: intranasal beta-estradiol
Comparison: placebo

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with placebo	Risk with Beta-estradiol intranasal				
Frequency of hot flushes Change from baseline Intranasal beta-estradiol 0.021 mg/day, 0.029 mg/day or 0.040 mg/day	The mean rate of hot flushes per day was lower in the beta-estradiol group (MD - 3.04, 95% CI -4.05 to -2.03)			458 (1 RCT)	⊕⊕⊕◯ moderate[1]	
Adverse effects Intranasal beta-estradiol 0.021 mg/day, 0.029 mg/day or 0.040 mg/day	171 per 1000	288 per 1000 (206 to 385)	OR 1.96 (1.26 to 3.03)	458 (1 RCT)	⊕⊕⊕◯ moderate[1]	

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the median risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

CI: Confidence interval; MD: mean difference; **OR:** Odds ratio; mg/d: milligrams per day

GRADE Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect

Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

1. Downgraded one level due to serious risk of attrition bias

Beta-estradiol versus conjugated equine estrogen for hot flushes

Population: women with hot flushes
Setting: community
Intervention: beta-estradiol patch
Comparison: conjugated equine estrogen (CEE)

Outcomes	Beta-estradiol versus CEE	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
Frequency of hot flushes Beta-estradiol patch 0.05 or 0.1 mg per day vs CEE 0.625 or 1.25 mg per day	Two studies with data unsuitable for analysis reported no evidence of a difference between the groups	711 (2 RCTs)	⊕⊕◯◯ low [1]	Results in graphics
Adverse events Beta-estradiol patch 0.05 or 0.1 mg/day vs CEE 0.625 or 1.25 mg/day	Two studies reported data unsuitable for analysis. One reported higher rates of breast pain and vaginal bleeding in the 0.1 mg/day beta-estradiol group than in the placebo group. There was no evidence of a difference between other groups	711 (2 RCTs)	⊕◯◯◯ very low [1,2]	

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.
OR: odds ratio; CI: Confidence interval; mg/d: milligrams per day

[1] Downgraded two levels for very serious risk of bias: studies at high or unclear risk of bias in most domains. Findings unreliable as no statistical data suitable for analysis
[2]Downgraded one level for serious imprecision, with low event rate

Oral beta-estradiol versus conjugated equine estrogen for hot flushes

Population: women with hot flushes
Setting: community
Intervention: oral beta-estradiol
Comparison: conjugated equine estrogen (CEE)

Outcomes	Beta-estradiol versus CEE	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
Frequency of hot flushes Oral beta-estradiol 1.0 and 2.0 mg/day vs CEE 0.625 or 1.25 mg/day	One study with data unsuitable for analysis reported no evidence of a difference between the groups	102 (1 study)	⊕◯◯◯ very low[1,2]	
Adverse events Oral beta-estradiol 1.0 or 2.0 mg/day vs CEE 0.625 or 1.25 mg/day	One study reported no evidence of a difference between the groups (OR 1.20, 95% CI 0.50 to 2.87)	103 (1 RCT)	⊕◯◯◯ very low[3]	

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.
OR: odds ratio; CI: Confidence interval; mg/d: milligrams per day

[1] Downgraded two levels for very serious risk of bias: studies at high or unclear risk of bias in most domains. Findings unreliable as no statistical data suitable for analysis
[2]Downgraded one level for serious imprecision, with small sample size
[3]Downgraded one level for very serious risk of bias (methods not adequately reported) and two levels for very serious imprecision (low event rate and small sample size)

DISCUSSION

Summary of main results

This review assessed the effectiveness and safety of bio-identical hormones (beta-estradiol E2) compared to placebo, or non bioidentical hormones for the relief of vasomotor symptoms during the menopausal transition.

According to the findings of this systematic review, treatment with oral and transdermal beta-estradiol is more effective in the relief of menopausal hot flushes than placebo. There was no good evidence of a difference in effectiveness between BHT and CEE but the evidence was insufficient to reach a definite conclusion. Even in extremely low doses, estrogen patches are more effective than placebo in decreasing the number of daily hot flushes in women experiencing at least seven hot flushes per day or 50 hot flushes per week at baseline. Our findings suggest that, regardless of the route of administration, low doses of BHT can reduce hot flushes in these women. An advantage of the transdermal route, compared to the oral route, is that it is available in many different forms and doses (Kopper 2009). In comparisons between BHT and placebo, adverse events were more common in the intervention arm. Breast tenderness (Archer 1992; de Vrijer 2000; Speroff 1996; von Holst 2000) and vaginal bleeding (De Aloysio 2000; de Vrijer 2000; Good 1999; Notelovitz 2000a; Wiklund 1993) were the most commonly reported adverse effects among estrogen users. Two trials reported one case of endometrial cancer each in a woman using beta-estradiol (de Vrijer 2000; Notelovitz 2000b). Bleeding and breast tenderness were more frequent among women using higher rather than lower doses of estrogen, regardless of the type of estrogen (Archer 1992). Adverse skin reactions were most common among women using transdermal forms of beta-estradiol or placebo (Bacchi-Modena 1997; De Aloysio 2000; Gordon 1995; Speroff 1996; Utian 1999; Wiklund 1993). The most frequent side-effects were breast tenderness, headache and skin reactions. The prevalence of these symptoms was related to the dosages and routes of administration (Steingold 1991). According to the most recent U.S. Food and Drug Administration (FDA) recommendation, hormonal therapy should be given at the lowest effective dose for the least amount of time necessary for women in the menopausal transition (NAMS 2012).

In comparisons between BHT and CEE, findings with regard to adverse effects were inconsistent and the quality of the evidence was too very low to reach any firm conclusions (Archer 1992; Good 1999; Gordon 1995)

Subgroup analyses of comparisons between BHT and placebo suggested that increased doses of BHT may be associated with increased effectiveness but also increased risk of adverse effects.

Overall completeness and applicability of evidence

This review identified 23 studies which evaluated bioidentical hormones for postmenopausal women with amenorrhoea and elevated FSH. Most of the trials included in this systematic review followed participants only up to 12 weeks. This is a limitation because these follow-up periods are too short to assess safety outcomes such as those included in the WHI study. One of the studies (De Vrijer 2000) noted that five women in the intervention grouped developed endometrial hyperplasia as no progestogens were administered during the trial treatment in order to avoid any effect of progestogens on the efficacy data. Studies with longer follow-ups which include women who still have their uterus would need to give participants progestogens or micronised progesterone, which is a bioidentical hormone used to prevent estrogen-related endometrial hyperplasia (NAMS 2012; Furness 2012). Another limitation of this systematic review is the lack of studies comparing beta-estradiol combined with progesterone. This precludes us from assessing if this type of prescription would alter the effects of beta-estradiol.

Although we did not restrict study eligibility by severity of vasomotor symptoms, the findings of this review apply largely to women with moderate to severe hot flushes, since most of the studies included only women with a minimum of three to seven hot flushes per day or 56 to 60 hot flushes per week of moderate to severe intensity.

The FDA 2003 recommends that to test the efficacy of treatments for vasomotor symptoms, the mean change in symptoms from baseline should be measured at four and 12 weeks. All studies included in this review followed this recommendation except Gordon 1995 which performed measurements at 11 weeks. As noted above, the limited follow-up in the included studies means that no data were reported about the safety of BHT with regard to long-term outcomes such as heart attack, stroke and breast cancer.

Quality of the evidence

Eight of 23 studies included in this review we deemed to be at high risk of bias in one or more domains. Limitations included very poor reporting of methods, attrition bias due to high or unbalanced dropouts between groups, and other potential sources of bias related to study design. Most studies were either funded by industry or the funding source was not reported. Many of the studies failed to report data suitable for analysis.

The overall quality of the evidence on the effect of beta-estradiol ranged from very low to moderate. Common limitations were risk of bias, inconsistency and imprecision. (Summary of findings for the main comparison; Summary of findings 2; Summary of findings 3; Summary of findings 5; Summary of findings 7).

Unfortunately a large number of studies were not suitable for pooling, in most cases due to skewed data associated with very high standard deviations, and this limited the power of our meta-analyses. However, the findings of studies with data unsuitable for pooling supported the findings of pooled analyses.

Potential biases in the review process

We made all efforts to identify relevant articles for this review by an extensive electronic search and careful evaluation of study reports for eligibility. Nevertheless, we cannot rule out the possibility that we may have lost some trials not accessible to our search. This review was limited because of the methodological heterogeneity of the included studies. The incompleteness of some of the reports and our difficulties in obtaining clarification of certain trial details or in resolving ambiguities in the reports may have contributed to some bias in their assessment. Duplicate study selection, data extraction and quality assessment of the included studies minimised the potential for additional bias beyond those detailed in the 'Risk of bias in included studies' tables.

Agreements and disagreements with other studies or reviews

These results are consistent with other reviews and meta-analyses of trials of transdermal beta-estradiol compared with placebo for treating menopausal hot flushes (Corbelli 2014; Heidi 2004). The intensity of symptoms was also significantly lower in all intervention groups compared to placebo. Two studies compared transdermal beta-estradiol gel versus placebo (Archer 2003; Archer 2012). In all sets of comparisons there was a significant difference favouring the intervention group compared with placebo. Compared with placebo, oral beta-estradiol produced statistically significant improvements in hot flush frequency or severity, or both.

AUTHORS' CONCLUSIONS

Implications for practice

There is low to moderate quality evidence that BHT in various forms and doses is more effective than placebo in decreasing the frequency of moderate to severe hot flushes in women in the menopausal transition period. There was low to moderate quality evidence of higher rates of adverse effects such as breast pain, skin reactions, vaginal bleeding and headache in the BHT group. There is some evidence to suggest that higher doses of BHT are more effective but also associated with a higher risk of adverse effects. Women with a uterus who are taking any form of estrogen require co-administration of a progestogen, as unopposed estrogen is associated with endometrial hyperplasia (Furness 2012).

There was no good evidence of a difference in effectiveness between BHT and CEE and findings with regard to adverse effects were inconsistent. The quality of the evidence was too low to reach any firm conclusions.

The main limitations in the quality of the evidence were study risk of bias (mainly due to poor reporting of methods), imprecision and lack of data suitable for analysis. No data are yet available about the safety of BHT with regard to long-term outcomes such as heart attack, stroke and breast cancer.

Implications for research

Studies that follow the guideline recommendations on hormone therapy for the relief of vasomotor symptoms in women during the menopausal transition (using the lowest effective dose through the best route of administration) are needed to confirm the few positive results that have been reported in randomised clinical trials to date. The review did not provide sufficient evidence to offer clinicians a firm conclusion as to the best dose and route of administration of bioidentical hormones (bioidentical estrogen alone or in combination with progesterone) for treating women with vasomotor symptoms. Data from trials evaluated in this review do not allow comparisons of adverse effects because they were reported in incomplete and nonstandardised ways. The symptom treatment trials reviewed herein enrolled small numbers of participants for short periods of time and were inadequately designed to capture important health outcomes reported by the Womens Health Initiative study (Rossouw 2002). Future trials could address these issues by providing a broader demographic sample of women, longer follow-up periods, larger numbers of participants, and more head-to-head comparisons of BHT (with or without progesterone) versus equine products, estradiol valerate, micronised estradiol, estropipate (piperazine estrone sulfate), ethinyl estradiol alone or in combination with various progestogens (medroxyprogesterone acetate, norethindrone, norethindrone acetate, drospirenone, dienogest) administered in cyclic or continuous regimens.

Research is needed to investigate the safety of BHT with regard to long-term outcomes such as heart attack, stroke and breast cancer.

ACKNOWLEDGEMENTS

We would like to thank the Cochrane Gynaecology and Fertility group Editorial Team, mainly Ms. Helen Nagels for their thoughtful comments and suggestions, and Information Specialist Marian Showell. We also want to acknowledge Cristiane R Macedo for assistance with developing the first search strategy and Evilásio Rodrigues Cortes for the valuable technical help with the Review Manager (RevMan) (RevMan 2014) programme.

REFERENCES

References to studies included in this review

Archer 1992 *{published data only}*
Archer DR, Fischer LA, Rich D, Schade GH, Schwartz S, Wittcoff H, et al. Estrace vs Premarin for treatment of menopausal symptoms: dosage comparison study. *Advances in Therapy* 1992;**9**:21–31.

Archer 2003 *{published data only}*
Archer DF. Percutaneous 17-estradiol gel for the treatment of vasomotor symptoms in postmenopausal women. *The Journal of The North American Menopause Society* 2003;**10**(6):516–21. [DOI: 10.1097/01

Archer 2012 *{published data only}*
Archer DF, Pickar JH, MacAllister DC, Warren MP. Transdermal estradiol gel for the treatment of symptomatic postmenopausal women. *Menopause* 2012;**19**(6):622–9.

Bacchi-Modena 1997 *{published data only}*
Bacchi-Modena A, Bacchi-Modena C, Bolis P, Campagnoli M, De Cicco R, Meschia G, et al. Efficacy and tolerability of Estraderm MX, a new estradiol matrix patch. *Maturitas* 1997;**27**:285–92.

Buster 2008 *{published data only}*
Buster JE, Koltun WD, Pascual ML, Day WW, Peterson C. Low-dose estradiol spray to treat vasomotor symptoms. *Obstetrics and Gynecology* 2008;**111**:1343-51. CENTRAL: CN–00640015]

Cohen 1999 *{published data only}*
Cohen L, Coxwell WL, Melchione T, Koltun W, Gibson E, Gupta N, et al. Low-dose 17-beta estradiol matrix transdermal system in the treatment of moderate-to-severe hot flushes in postmenopausal women. *Current Therapeutic Research - Clinical and Experimental* 1999;**60**(10):534–47.

De Aloysio 2000 *{published data only}*
De Aloysio D, Rovati LC, Giacovelli G, Setnikar I, Bottiglioni F. Efficacy on climacteric symptoms and safety of low dose estradiol transdermal matrix patches. *Azneimittelforschung* 2000;**50**:293–300.

De Vrijer 2000 *{published data only}*
De Vrijer B, Snijders MP, Troostwijk AL, Thé RJ, Iding Da Freise JM, Smit H, et al. Efficacy and tolerability of a new estradiol delivering matrix patch (Estraderm MX) in postmenopausal women. *Maturitas* 2000;**34**:47–55.

Good 1999 *{published data only}*
Good WR, John VA, Ramirez M, Higgins JE. Comparison of Alora estradiol matrix transdermal delivery system with oral conjugated equine estrogen therapy in relieving menopausal symptoms. *Climacteric* 1999;**2**:29–36.

Gordon 1995 *{published data only}*
Gordon SF, Thompson KA, Ruoff GE, Imig JR, Lane PJ, Schwenker CE. Efficacy and safety of a seven-day, transdermal estradiol drug-delivery system: comparison with conjugated estrogens and placebo. *International Journal of Fertility and Menopausal Studies* 1995;**40**:126–34.

Haines 2009 *{published data only}*
Haines C, Yu SL, Hiemeyer F, Schaefers M. Micro-dose transdermal estradiol for relief of hot flushes in postmenopausal Asian women: a randomized controlled trial. *Climacteric* 2009;**12**:419-26.

Honjo 2009 *{published data only}*
Honjo H, Taketani Y. Low-dose estradiol for climacteric symptoms in Japanese women: a randomized, controlled trial. *Climacteric* 2009;**12**:319-28. [DOI: 10.1080/13697130802657888

Nielsen 2006 *{published data only}*
Nielsen TF, Ravn P, Pitkin J, Christiansen C. Pulsed estrogen therapy improves postmenopausal quality of life: a 2-year placebo-controlled study. *Maturitas* 2006;**53**:184–90. [DOI: 10.1016

Notelovitz 2000a *{published data only}*
Notelovitz M, Lenihan JP, McDermott M, Kerber IJ, Nanavati N, Arce J. Initial 17beta-estradiol dose for treating vasomotor symptoms. *Menopause* 2000;**7**:310–17.

Notelovitz 2000b *{published data only}*
Notelovitz M, Mattox JH. Suppression of vasomotor and vulvovaginal symptoms with continuous oral 17beta-estradiol. *Menopause* 2000;**7**:310–317.

Rovati 2000 *{published data only}*
Rovati LC, Setnikar I, Genazzani AR. Dose-response efficacy of a new estradiol transdermal matrix patch for 7-day application: a randomized, double-blind, placebo-controlled study. *Menopause: The Journal of The North American Menopause Society* 2002;**22**(1):114–21. [DOI: 10.1097

Simon 2006 *{published data only}*
Simon JA. Estradiol in micellar nanoparticles: the efficacy and safety of a novel transdermal drug-delivery technology in the management of moderate to severe vasomotor symptoms. *Menopause: The Journal of The North American Menopause Society* 2006;**13**(2):222–31. DOI: 10.1097/01.gme.0000174096.56652.4f]

Simon 2007 *{published data only}*
Simon JA, Bouchard C, Waldbaum A, Utian W, Zborowski J, Snabes MC. Low dose of transdermal estradiol gel for treatment of symptomatic postmenopausal women: a randomized controlled trial. *Obstetrics and Gynecology* 2007;**109**(3):588–96.

Speroff 1996a *{published data only}*
Speroff L, Whitcomb RW, Kempfert NJ, Boyd RA, Paulissen JB, Rowan JP. Efficacy and local tolerance of a low-dose, 7-day matrix estradiol transdermal system in the treatment of menopausal vasomotor symptoms. *Obstetrics and Gynecology* 1996;**88**(4):587–92.

Speroff 1996b *{published data only}*
Speroff L, Whitcomb RW, Kempfert NJ, Boyd RA, Paulissen JB, Rowan JP. Efficacy and local tolerance of a low-dose, 7-day matrix estradiol transdermal system in the

treatment of menopausal vasomotor symptoms. *Obstetrics and Gynecology* 1996;**88**(4):587–92.

Utian 1999 ***{published data only}***
Utian WH, Burry KA, Archer DF. Efficacy and safety of low, standard, and high dosages of an estradiol transdermal system (Esclim) compared with placebo on vasomotor symptoms in highly symptomatic menopausal patients. *American Journal of Obstetrics and Gynecology* 1999;**181**: 71–9.

von Holst 2000 ***{published data only}***
von Holst T, Sallbach B. Efficacy and tolerability of a new 7-day transdermal estradiol patch versus placebo in hysterectomized women with postmenopausal complaints. *Maturitas* 2000;**34**:143–53.

Wiklund 1993 ***{published data only}***
Wiklund I, Karlberg J, Mattsson LA. Quality of life of postmenopausal women on a regimen of transdermal estradiol therapy: a double-blind placebo-controlled study. *American Journal of Obstetrics and Gynecology* 1993;**168**: 824–30.

References to studies excluded from this review

Akhila 2006 ***{published data only}***
Akhila V, Pratapkumar. A comparison of transdermal and oral HRT for menopausal symptom control. *International Journal of Fertility and Women's Medicine* 2006;**51**(2):64–9.

Bachmann 2003 ***{published data only}***
Bachmann GA. Estrogen no matter how delivered relieves postmenopausal vasomotor symptoms. *Menopause* 2003;**10** (6):494–6.

Bachmann 2007 ***{published data only}***
Bachmann GA, Schaefers M, Uddin A, Utian WH. Lowest effective transdermal 17-estradiol dose for relief of hot flushes in postmenopausal women. *Obstetrics and Gynecology* 2007;**110**(4):771-9.

Bachmann 2008 ***{published data only}***
Bachmann GA, Schaefers M, Uddin A, Utian WH. What's the lowest effective estrogen dose for hot flushes?. *Journal of Family Practice* 2008;**57**(1):9.

Ben-Chetrit 2005 ***{published data only}***
Ben-Chetrit A, Hochner-Celnikier D, Lindenberg T, Zacut D, Shimonovitz S, Gelber H, et al. Vaginal ring delivering estradiol and progesterone: a possible alternative to relieve climacteric symptoms. *Israeli Medical Association Journal* 2005;**7**:302–6.

Carranza-Lira 2006 ***{published data only}***
Carranza-Lira S, Gooch AL, Velasco-Díaz G, Solano J, Arzola-Paniagua A. Low and ultra low-dose estrogen therapy for climacteric symptom control - preliminary report. *International Journal of Fertility and Women's Medicine* 2006; **51**(4):171–5.

Castelo-Branco 2010 ***{published data only}***
Castelo-Branco C, Coloma JL. The role of intranasal estradiol spray in the management of moderate to severe vasomotor symptoms in menopausal women. *Gynecological Endocrinology* 2010;**26**(1):23–9. [DOI: 10.3109/ 09513590903159698

Chung 1996 ***{published data only}***
Chung TKH, Yip SK, Lam P, Chang AMZ, Haines CJ. A randomized, double-blind, placebo-controlled, crossover study on the effect of oral oestradiol on acute menopausal symptoms. *Maturitas* 1996;**25**:115–23.

Conaway 2011 ***{published data only}***
Conaway E. Bioidentical hormones: an evidence-based review for primary care providers. *Journal of the American Osteopathic Association* 2011;**111**(3):153–64.

Cortes-Bonilla 2015 ***{published data only}***
Cortés-Bonilla M, Bernardo-Escudero R, Alonso-Campero R, Francisco-Doce MT, Hernández-Valencia M, Celis-González C, et al. Treatment of menopausal symptoms with three low-dose continuous sequential 17β-estradiol/ progesterone parenteral monthly formulations using novel non-polymeric microsphere technology. *Gynecological Endocrinology* 2015;**10**:1–8.

Darj 1991 ***{published data only}***
Darj E, Nilsson S, Axelsson O, Hellberg D. Clinical and endometrial effects of oestradiol and progesterone in post-menopausal women. *Maturitas* 1991;**13**(2):109–15.

Diem 2006 ***{published data only}***
Diem S, Grady D, Quan J, Vittinghoff E, Wallace R, Hanes V, et al. Effects of ultra low-dose transdermal estradiol on postmenopausal symptoms in women aged 60 to 80 years. *Menopause* 2006;**13**(1):130–8. [DOI: 10.1097/ 01.gme.0000192439.82491.24

Ettinger 2007 ***{published data only}***
Ettinger B. Rationale for use of lower estrogen doses for postmenopausal hormone therapy. *Maturitas* 2007;**57**(1): 81–4.

Files 2011 ***{published data only}***
Files JA, Ko MG, Pruthi S. Bioidentical hormone therapy. *Mayo Clinic Proceedings* 2011;**86**(7):673.

Formby 2011 ***{published data only}***
Formby B, Schmidt F. Efficacy of biorhythmic transdermal combined hormone treatment in relieving climacteric symptoms: a pilot study. *International Journal of General Medicine* 2011;**4**:159–63.

Ganz 2002 ***{published data only}***
Ganz P. Vasomotor and vascular effects of hormone replacement therapy. *American Journal of Cardiology* 2002; **90**:F11–16. PUBMED: 12106634]

Gass 2004 ***{published data only}***
Gass MS, Rebar RW, Cuffie-Jackson C, Cedars MI, Loboe RA, Shoupe D, et al. A short study in the treatment of hot flashes with buccal administration of 17- estradiol. *Maturitas* 2004;**49**:140–7.

Hedrick 2009 ***{published data only}***
Hedrick RE, Ackerman RT, Koltun WD, Halvorsen MB, Lambrecht LJ. Transdermal estradiol gel 0.1% for the treatment of vasomotor symptoms in postmenopausal women. *The Journal of The North American*

Menopause Society 2009;**16**(1):132–40. [DOI: 10.1097/gme.0b013e31817d5372

Iftikhar 2011 ***{published data only}***
Iftikhar S, Shuster LT, Johnson RE, Jenkins SM, Wahner-Roedler DL. Use of bioidentical compounded hormones for menopausal concerns: cross-sectional survey in an academic menopause center. *Journal of Women's Health* 2011;**20**(4):559–65. [DOI: 10.1089

Jensen 1987 ***{published data only}***
Jensen PB. Climacteric symptoms after oral and percutaneous hormone replacement therapy. *Maturitas* 1987;**9**:207–15.

Lacut 2004 ***{published data only}***
Lacut K, et al. Effects of oral and transdermal 17 beta-estradiol combined with progesterone on homocysteine metabolism in postmenopausal women: a randomised placebo-controlled trial. *Atherosclerosis* 2004;**174**(1):173-80. PUBMED: 15135267]

Lindh 2004 ***{published data only}***
Lindh-Åstrand L, Nedstrand E, Wyon Y, Hammar M. Vasomotor symptoms and quality of life in previously sedentary postmenopausal women randomised to physical activity or estrogen therapy. *Maturitas* 2003;**48**:97–105.

Lopes 2000 ***{published data only}***
Lopes P, Merkus HM, Nauman J, Bruschi F, Foidart JM, Calaf J. Randomized comparison of intranasal and transdermal estradiol. *Obstetrics and Gynecology* 2000;**96**(6):906–12.

Lubbert 1997 ***{published data only}***
Lübbert H, Nauert C. Continuous versus cyclical transdermal estrogen replacement therapy in postmenopausal women: influence on climacteric symptoms, body weight and bleeding pattern. *Maturitas* 1997;**28**:117-25.

MacLennan 2009 ***{published data only}***
MacLennan AH. Evidence-based review of therapies at the menopause. *International Journal of Evidence Based Healthcare* 2009;**7**:112–23. [DOI: 10.1111

Marslew 1991 ***{published data only}***
Marslew U, Riis B, Christiansen C. Progestogens: therapeutic and adverse effects in early post-menopausal women. *Maturitas* 1991;**13**(1):7–16.

Marslew 1994 ***{published data only}***
Marslew U, Munk-Nielsen N, Nilas L, Riis BJ, Christiansen C. Bleeding pattern and climacteric symptoms during different sequential combined HRT regimens in current use. *Maturitas* 1994;**19**(3):225–37.

Mather 2000 ***{published data only}***
Mather KJ, Norman EG, Prior JC, Elliott TG. Preserved forearm endothelial responses with acute exposure to progesterone: a randomized cross-over trial of 17-b estradiol, progesterone, and 17-b estradiol with progesterone in healthy menopausal women. *The Journal of Clinical Endocrinology and Metabolism* 2000;**85**(12):4644-49.

Mirkin 2015 ***{published data only}***
Mirkin S, Amadio JM, Bernick BA, Pickar JH, Archer DF. 17(beta)-Estradiol and natural progesterone for menopausal hormone therapy: REPLENISH phase 3 study design of a combination capsule and evidence review. *Maturitas* 2015;**81**(1):28–35.

Mizumuna 2011 ***{published data only}***
Mizunuma H. Clinical usefulness of a low-dose maintenance therapy with transdermal estradiol gel in Japanese women with estrogen deficiency symptoms. *Climacteric* 2011;**14**:581-89. CENTRAL: CN–00806509]

Odabasi 2007 ***{published data only}***
Odaba i AR, Yüksel H, Demircan SS, Kaçar DF, Çulhaci N, Özkara EE. A prospective randomized comparative study of the effects of intranasal and transdermal 17ß estradiol on postmenopausal symptoms and vaginal cytology. *Journal of Postgraduate Medicine* 2007;**53**:221–7.

Panay 2007 ***{published data only}***
Panay N, Ylikorkala O, Archer DF, Gut R, Lang E. Ultra-low-dose estradiol and norethisterone acetate: effective menopausal symptom relief. *Climacteric* 2007;**10**:120–31.

Pélissier 2001 ***{published data only}***
Pélissier C, Maroni M, Yaneva H, Brin S, Peltier-Pujol F, Jondet M. Chlormadinone acetate versus micronized progesterone in the sequential combined hormone replacement therapy of the menopause. *Maturitas* 2001;**40**(1):85–94.

Rosano 2000 ***{published data only}***
Rosano GMC, Webb CM, Chierchia S, Morgani GL, Gabraele M, Sarrel PM, et al. Natural progesterone, but not medroxyprogesterone acetate, enhances the beneficial effect of estrogen on exercise-induced myocardial ischemia in postmenopausal women. *Journal of the American College of Cardiology* 2000;**36**(7):2154–9.

Ryan 2001 ***{published data only}***
Ryan N, Rosner A. Quality of life and costs associated with micronized progesterone and medroxyprogesterone acetate in hormone replacement therapy for nonhysterectomized, postmenopausal women. *Clinical Therapeutics* 2001;**23**(7):1099-1115. [DOI: 10.1016

Serfaty 2003 ***{published data only}***
Serfaty D, Reilhac P, Eschwege E, Ringa V, Blin P, Nandeuil A, et al. Compliance with hormone replacement therapy in menopausal women: results of a two-year prospective French study comparing transdermal treatment with fixed oral combination therapy [Observance du traitement hormonal substitutif de la ménopause: résultats d'une étude prospective française de deux ans, comparantun traitement transdermique et une association fixe par voie orale]. *Gynécologie Obstétrique and Fertilité* 2003;**31**:525-33.

Sitruk 2007 ***{published data only}***
Sitruk-Ware R. Routes of delivery for progesterone and progestins. *Maturitas* 2007;**57**:77-80.

Siyam 2013 *{published data only}*
Siyam T, Yuksel N. Beliefs about bioidentical hormone therapy: a cross-sectional survey of pharmacists. *Maturitas* 2013;**74**:196–202.

Skarsgard 2000 *{published data only}*
Skarsgard C, Berg GE, Ekblad S, Wiklund I, Hammar ML. Effects of estrogen therapy on well-being in postmenopausal women without vasomotor complaints. *Maturitas* 2000;**36**: 123–30.

Somunkiran 2007 *{published data only}*
Somunkiran A, Erel CT, Demirci F, Senturk ML. The effect of tibolone versus 17-estradiol on climacteric symptoms in women with surgical menopause: a randomized, cross-over study. *Maturitas* 2007;**56**:61–8.

Sood 2011 *{published data only}*
Sood R, Shuster L, Smith R, Vincent A, Jatoi A. Counseling postmenopausal women about bioidentical hormones: ten discussion points for practicing physicians. *Journal of the American Board of Family Medicine* 2011;**24**(2):202-10.

Sood 2013 *{published data only}*
Sood R, Warndahl RA, Schroeder DR, Singhd R J, Rhodes DJ, Wahner-Roedlerf D, et al. Bioidentical compounded hormones: a pharmacokinetic evaluation in a randomized clinical trial. *Maturitas* 2013;**74**:375- 82.

Studd 1999 *{published data only}*
Studd J, Pornel B, Marton I, Bringer J, Varin C, Tsouderos Y, et al. Efficacy and acceptability of intranasal 17 - oestradiol for menopausal symptoms: randomised dose-response study. *Lancet* 1999;**353**:1574–8.

Suvanto-luukkonen 1997 *{published data only}*
Suvanto-Luukkonen E, Sundstrijm E, Jorma Penttinen J, Läärä E, Pramila S, Kauppila A. Percutaneous estradiol gel with an intrauterine levonorgestrel releasing device or natural progesterone in hormone replacement therapy. *Maturitas* 1997;**26**:211–17.

Vartiainen 1993 *{published data only}*
Vartiainen J, Wahlstriim T, Nilsson CG. Effects and acceptability of a new 17B-oestradiol releasing vaginal ring in the treatment of postmenopausal complaints. *Maturitas* 1993;**17**:129–37.

Veerus 2013 *{published data only}*
Veerus P, Hovi SL, Sevón T, Hunter M, Hemminki E. The effect of hormone therapy on women's quality of life in the first year of the Estonian Postmenopausal Hormone Therapy trial. *Menopause* 2013;**20**(3):291–8. [DOI: 10.1097/GME.0b013e31826ce3ed.

Whelan 2013 *{published data only}*
Whelan AM, Jurgens TM, Trinacty M. Bioidentical progesterone cream for menopause-related vasomotor symptoms: is it effective?. *The Annals of Pharmacotherapy* 2013;**47**:112–6. [DOI: 10.1345

Wihlbacka 2005 *{published data only}*
Wihlbäck A, Nyberg S, Bäckström T, Bixo M, Sundstrom-Poromaa I. Estradiol and the addition of progesterone increase the sensitivity to a neurosteroid in postmenopausal women. *Psychoneuroendocrinology* 2005;**30**(1):38–50.

Wolfe 1994 *{published data only}*
Wolfe BM, Plunkettb ER. Early effects of continuous low-dosage dl-norgestrel administered alone or with estrogen. *Maturitas* 1994;**18**:207–19.

Yesildaglar 2004 *{published data only}*
Yesildaglar N, Erkaya S, Uygur D, Göl K, Bingöl B, Günenç Z. Efficacy of pulsed estrogen therapy in relatively younger patients with surgically induced menopause. *Human Reproduction* 2004;**19**(1):210–3.

Additional references

Adams 2001
Adams C, Cannell S. Women's beliefs about "natural" hormones and natural hormone replacement therapy. *Menopause* 2001;**8**:433–40.

Cagnacci 2002
Cagnacci A, Arangino S, Tuveri F, Paoletti AM, Volpe A. Regulation of the 24h body temperature rhythm of women in luteal phase: role of gonadal steroids and prostaglandins. *Chronobiology International Journal* 2002;**19**(4):721–30.

Corbelli 2014
Corbelli J, Shaikh N, Wessel C, Hess R. Low-dose transdermal estradiol for vasomotor symptoms: a systematic review. *Menopause: The Journal of The North American Menopause Society* 2014;**22**(1):114–21. [DOI: 10.1097/gme.0000000000000258

Corson 1993
Corson SL. A decade of experience with transdermal estrogen replacement therapy: overview of key pharmacologic and clinical findings. *International Journal of Fertility* 1993;**38**: 79–81.

Deeks 2011
Deeks JJ, Higgins JPT, Altman DG (editors). Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Drisko 2000
Drisko JA. Natural isomolecular hormone replacement: an evidence-based medicine approach. *International Journal of Pharmaceutical Compounding* 2000;**4**:414–20.

EMEA 2005
European Medicines Agency. Guideline on clinical investigation of medicinal products for hormone replacement therapy of oestrogen deficiency symptoms in postmenopausal women. ema.europa.eu/ema/ Accessed September 2012.

ES 2006
Endocrine Society. Bioidentical hormones: position statement. endo-society.org/advocacy/policy/upload/BH_position_Statement_final_10_25_06_w_Header.pdf Accessed May 20, 2012.

FDA 2003
US Department of Health and Human Services. Guidance for industry: estrogen and estrogen/progestin drug

products to treat vasomotor symptoms and vulvar and vaginal atrophy symptoms-recommendations for clinical evaluation. fda.gov/cder/guidance/index.htm January 2003. [fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071643.pdf(accessed 2014 july)]

Freedman 2001

Freedman RR. Physiology of hot fl ashes. *American Journal of Human Biology* 2001;**13**:453.

Freedman 2006

Freedman RR, Benton MD, Genik II RJ, Graydon FX. Cortical activation during menopausal hot flashes. *Fertility and Sterility* 2006;**85**:674.

Furness 2012

Furness S, Roberts H, Marjoribanks J, Lethaby A. Hormone therapy in postmenopausal women and risk of endometrial hyperplasia. *Cochrane Database of Systematic Reviews* 2012, Issue 8. [DOI: 10.1002/14651858; CD000402

GRADEpro GDT 2015 [Computer program]

McMaster University (developed by Evidence Prime, Inc). Available from gradepro.org. GRADEpro Guideline Development Tool. McMaster University (developed by Evidence Prime, Inc). Available from gradepro.org, 2015.

Harlow 2012

Harlow SD, Gass M, Hall JE, Lobo R, Maki P, Rebar RW, et al. Executive summary of the Stages of Reproductive Aging Workshop D10: addressing the unfinished agenda of staging reproductive aging. *Journal of Clinical Endocrinology and Metabolism* 2012;**97**(4):1159–68.

Heidi 2004

Heidi DN. Commonly used types of postmenopausal estrogen for treatment of hot flashes. *JAMA* 2004;**219**(13): 1610–20. [DOI: 10.1001/jama.291.13.1610

Hersh 2003

Hersh AL, Stefanick ML, Stafford RS. National use of postmenopausal hormone therapy: annual trends and response to recent evidence. *JAMA* 2004;**291**:47–53.

Higgins 2003

Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327**: 557–60.

Higgins 2011a

Higgins JPT, Altman DG, Sterne JAC (editors). Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Higgins 2011b

Higgins JPT, Deeks JJ, Altman DG (editors). Chapter 16: Special topics in statistics. In: Higgins JPT, Green S (editors),*Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Hilditch 1996

Hilditch JR, Lewis J, Peter A, Van Maris B, Ross A, Franssen E, et al. A menopause-specific quality of life questionnaire: development and psychometric properties. *Maturitas* 1996; **24**(3):161–75.

Hunter 1992

Hunter MS. Women's Health Questionnaire: a measure of mid-aged women's perceptions of their emotional and physical health. *Psychology and Health* 1992;**7**(1):45–54.

Kaufert 1998

Kaufert P, Boggs PP, Ettinger B, Woods NF, Utian WH. Women and menopause: beliefs, attitudes, and behaviours. *Menopause* 1998;**5**:197.

Kopper 2009

Kopper NW, Gudeman J, Thompson DJ. Transdermal hormone therapy in postmenopausal women: a review of metabolic effects and drug delivery technologies. *Journal of Drug Design Development and Therapy* 2009;**6**(2):193–202. [PUBMED: 19920906]

Kronnenberg 1992

Kronnenberg F, Barnard RM. Modulation of menopausal hot flashes by ambient temperature. *Journal of Thermal Biology* 1992;**17**:43.

Lethaby 2013

Lethaby A, Marjoribanks J, Kronenberg F, Roberts H, Eden J, Brown J. Phytoestrogens for menopausal vasomotor symptoms. *The Cochrane Library* 2013;**12**:CD001395. [DOI: 10.1002/14651858

NAMS 2012

North American Menopause Society. Hormone therapy position statement of The North American Menopause Society. *The Journal of The North American Menopause Society* 2012;**19**(3):257–71.

Nelson 2006

Nelson HD, Vesco KK, Haney E, Fu R, Nedrow A, Miller J, et al. Nonhormonal therapies for menopausal hot flashes: systematic review and meta-analysis. *JAMA* 2006;**295**(17): 2057–69.

Nielsen 2004

Nielsen TF, Ravn P, Bagger YZ, Warming L, Christiansen C. Pulsed estrogen therapy in prevention of postmenopausal osteoporosis. A 2-year randomized, double blind, placebo-controlled study. *Osteoporosis International* 2004;**15**(2): 168–74.

RevMan 2014 [Computer program]

The Nordic Cochrane Centre, Cochrane. Review Manager (RevMan). Version 5.3. Copenhagen: The Nordic Cochrane Centre, Cochrane, 2014.

Rossouw 2002

Rossouw JE, Anderson GL, Prentice RL, LaCroix AZ, Kooperberg C, Stefanick ML, et al. Risks and benefits of estrogen plus progestin for healthy postmenopausal women. *JAMA* 2002;**288**:321–33.

Rossouw 2013
Rossouw JE, Manson JE, Kaunitz AM, Anderson GL. Lessons learned from the Women's Health Initiative trials of menopausal hormone therapy. *Obstetrics and Gynecology* 2013;**121**(1):172–6.

Seidl 1998
Seidl MM, Stewart DE. Alternative treatments for menopausal symptoms: qualitative study of women's experiences. *Canadian Family Physician* 1998;**44**:1271–6.

Shifren 2007
Shifren JL, Desindes S, McIlwain M, Doros G, Mazer NA. A randomised, open-label, crossover study comparing the effects of oral versus transdermal estrogens therapy on serum androgens, thyroid hormones, and adrenal hormones in naturally menopausal women. *Menopause* 2007;**14**:985.

Shifren 2010
Shifren JL, Schiff I. Role of hormone therapy in the management of menopause. *Obstetrics and Gynecology* 2010;**115**(4):839–55.

Speroff 2011
Speroff L, Glass RH, Kase NG, editors. Clinical Gynecologic Endocrinology and Infertility. *Clinical Gynecologic Endocrinology and Infertility*. 8th Edition. Philadelphia, PA: Lippincott Williams & Wilkins, 2011.

Steingold 1991
Steingold KA, Matt DW, DeZiegler D, Sealey JE, Fratkin M, Reznikov S. Comparison of transdermal to oral estradiol administration on hormonal and hepatic parameters in women with premature ovarian failure. *Journal of Clinical Endocrinology and Metabolism* 1991;**73**:275–80.

Sturdee 2011
Sturdee DW, Pines A, on behalf of the International Menopause Society Writing Group. Updated IMS recommendations on postmenopausal hormone therapy and preventive strategies for midlife health. *Climacteric* 2011;**14**:202–20.

Tartaryn 1981
Tartaryn IV, Lomax P, Meldrum DR, Bajorek JG, CHesorek W, Judd HL. Objective techniques for the assessment of menopausal hot flushes. *Obstetrics and Gynecology* 1981;**57**: 340–4.

Wilkin 1981
Wilkin JR. Flushing reactions: consequences and mechanisms. *Annals of Internal Medicine* 1981;**95**:468.

References to other published versions of this review

Gaudard 2013
Gaudard AMIS, da Silva EMK, Silva de Souza S, Torloni MR, Macedo CR. Bioidentical hormones for women with hot flushes. *Cochrane Database of Systematic Reviews* 2013, Issue 2. [DOI: 10.1002/14651858.CD010407

* *Indicates the major publication for the study*

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Archer 1992

Methods	Multicenter, double-blind, randomised, placebo-controlled trial Number of centres: 7 USA Number of women randomised: 128 Number of women analysed: 100 Withdrawals and losses to follow-up: n = 21.9 because of adverse events Duration of trial: 12 weeks Conflict of interest: not stated
Participants	Menopausal status: peri- and post-menopausal Age: mean 50.6 years, SD 5.9 Ethnicity: not stated Inclusion criteria: healthy, natural or surgically post-menopausal women, FSH > 40 IU/mL and serum oestradiol < 30 pg/mL (surgical menopause only), age 40-60 years (natural menopause only), moderate severe vasomotor symptoms (> 5/day of moderate to severe intensity) Exclusion criteria: significant past or present illness, genitourinary symptoms, psychological symptoms, gastrointestinal conditions, chronic headaches, any contraindications to beta-estradiol usage, HT within 3 months, concomitant medications that may affect study parameters, alcohol or drug abuse, laboratory abnormalities, Pap smear dysplasia, endometrial hyperplasia Confirmation of ovarian failure: FSH > 40 IU/mL and serum oestradiol < 30 pg/mL (surgical menopause only) Baseline equality: not reported for vasomotor symptoms. Groups equal for age, age at menopause, weight, history of HRT and history of abnormal Pap smear Baseline Symptoms: all women had moderate-severe vasomotor symptoms at baseline (inclusion criteria)
Interventions	Oral micronised beta-estradiol 1 or 2 mg/d versus conjugated equine estrogen (CEE) 0.625 or 1.25 mg/d or placebo
Outcomes	• Mean daily change from baseline of frequency and severity of moderate to severe hot flushes • Adverse events whether reported by the patient
Notes	Data unsuitable for analysis: reported in an additional table

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not stated
Allocation concealment (selection bias)	Unclear risk	Not reported

Archer 1992 (*Continued*)

Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Reported "double-blind". No further details
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not reported
Incomplete outcome data (attrition bias) All outcomes	High risk	Losses: 28/128 = 22%. Intention-to-treat analysis: no
Selective reporting (reporting bias)	Low risk	The study protocol was not available but it is clear that the published reports included all expected outcomes, including those that were pre-specified
Other bias	Low risk	The study appears to be free of other sources of bias

Archer 2003

Methods	Multicentre, double-blind, randomised, placebo-controlled trial Number of centres: 11 in USA Number of women randomised: 221 Number of women analysed: 196 Statistical analysis: baseline characteristics were compared among treatment groups with analysis of variance for continuous variables and the Cochrane-Mantel-Haenszel test and Fisher exact test for categorical variables. The primary efficacy variable was analysed using analysis of covariance and Dunnett's test for multiple comparisons. The incidence of adverse events was compared between active gel dose groups and the placebo group using Pearson X 2 test and a multiple comparisons procedure. All statistical tests were two-sided with an alpha of 0.049. Where multiple comparisons were used, the overall alpha level was 0.049 Withdrawals and losses to follow-up: n = 25. Only 4 of these women (1.8%) withdrew because of adverse events Duration of trial: 12 weeks Conflict of interest: none declared
Participants	Inclusion criteria: healthy postmenopausal women who were amenorrhoeic either naturally (for at least 6 months) or surgically (with bilateral oophorectomy), had a serum estradiol concentration no greater than 20 pg/mL, a serum FSH level of at least 40 mIU/mL, and had a minimum of seven hot flushes per day or 60 hot flushes per week of moderate-to-severe intensity (i.e. that interfered with daily activities, disrupted sleep, or were associated with perspiration) Exclusion criteria: treatment with oral, transdermal, or vaginal steroid hormones (estrogens, progestogens, androgens, or corticosteroids) at any time during the 8 weeks before the first screening visit or any of the following: allergy to estradiol; reactions to transdermally administered medications; beta-estradiol-dependent neoplasia; vascular disease,

Archer 2003 (*Continued*)

	thrombotic disorders, angina, active hepatic or gallbladder disease during the prior 6 months; or undiagnosed vaginal bleeding during the prior 6 months
Interventions	221 women were randomised to one of three gel groups: a) 1.25 g/day beta-estradiol gel containing 0.75 mg of estradiol (n = 75), b) 2.5 g/day beta-estradiol gel containing 1.5 mg estradiol (n = 73), or c) placebo gel (n = 73). The placebo gel was packaged to have an identical appearance to the beta-estradiol gel. The gel was applied once daily on the arm using a dispenser
Outcomes	• Mean daily change from baseline of frequency and severity of moderate to severe hot flushes • Diaries reviewed at weeks 4, 8, and 12 for incidence of adverse events
Notes	Treatment groups were comparable at baseline, except for mean weight which was significantly higher in the 1.25 g/day beta-estradiol gel group Data for effectiveness skewed; findings reported in an additional table

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not reported
Allocation concealment (selection bias)	Unclear risk	Not reported. Authors only reported that "Eligible women were randomly assigned to one of three treatment groups..."
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Authors reported that "The placebo gel was packaged to have an identical appearance to the 17-estradiol gel"
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	No details provided
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Losses 25/221 (11.3%). No ITT analysis
Selective reporting (reporting bias)	Low risk	The study protocol is not available but it is clear that the published reports include all expected outcomes, including those that were pre-specified
Other bias	Low risk	No other source of potential bias detected

Archer 2012

Methods	Multicentre, double-blind, randomised, placebo-controlled trial, phase 4 study Number of centres: 78 sites in the USA Number of women randomised: 351 Number of women analysed: 277 Statistical analysis: a sample size of 110 per group was estimated for the phase 4 study based on (1) a SD of 4.5 and a change of 2 for average daily moderate to severe hot flushes, (2) a SD of 0.72 and a change of 0.3 in hot flush severity score, and (3) a dropout rate of 12%, which would provide 80% power to reject the null hypothesis at weeks 4 and 12. A two-way analysis of covariance with fixed factors of treatment and centre and baseline value as a covariate was used to determine significant differences Withdrawals and losses to follow-up: n = 74 (21,1%). 12 withdrawn because of adverse events Duration of trial: 12 weeks Conflict of interest: none declared
Participants	Inclusion criteria: women who experienced at least an average of 7 moderate to severe hot flushes per day or at least 50 moderate to severe hot flushes per week during the 2-week screening period. Participants had serum estradiol concentrations of 20 pg/mL or less and FSH levels of 40 mIU/mL or higher in both studies at screening Exclusion criteria: known hypersensitivity to estrogens or contraindications to estrogen therapy; endometrial biopsy at study entry interpreted as hyperplasia; abnormal Papanicolaou test findings; body weight exceeding more than130% of the ideal range for weight and height or BMI of 39 kg/m2 or higher; uncontrolled hypertension; history of clinically significant migraines, cerebrovascular disease, thromboembolic events, or coronary heart disease or events; diabetes; liver disease; undiagnosed vaginal bleeding; malignancy of the cervix, uterus, adrenal glands, pituitary glands, breast, or ovaries; or melanoma or other skin cancers at an advanced stage. Treatments not permitted in the phase included a current intrauterine device; oral, transdermal, or vaginal steroid hormones (estrogens, progestogens, androgens, or natural remedies with hormonal effects); or selective estrogen receptor modulators within the last 30 days of screening; intrauterine progestin implant within 60 days; or oral contraceptives, gonadotropins, anti-estrogens, chronic high doses of corticosteroids, or thyroid medication not yet stabilised within 90 days
Interventions	0.375 mg beta-estradiol (1.25 g estradiol gel 0.03%), the expected no-effect dose of 0.27 mg estradiol (0.9 g estradiol gel 0.03%), versus placebo gel
Outcomes	• Mean daily change from baseline of frequency and severity of moderate to severe hot flushes • Adverse events were recorded at each study visit
Notes	The objective of this report was to compare the efficacy and safety of two doses (1.5 and 0.75 mg beta-estradiol) of a transdermal estradiol gel from a phase 3 study (Archer 2003) and two lower doses (0.375 and 0.27 mg beta-estradiol) from a phase 4 study (Archer 2012), to determine the lowest practical dose for the treatment of postmenopausal symptom and severity of hot flushes and vaginal cytology (vaginal maturation index) SDs not reported but when imputed, data for effectiveness were skewed; findings reported in an additional table
Risk of bias	

Archer 2012 (*Continued*)

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Authors only reported that "Eligible women were randomised.."
Allocation concealment (selection bias)	Unclear risk	Not reported
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Authors only reported "...double-blind... "No further details
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Authors only reported "...double-blind... "No further details
Incomplete outcome data (attrition bias) All outcomes	High risk	Losses 74/351 (21.1%). ITT analysis
Selective reporting (reporting bias)	Low risk	The study protocol is not available but it is clear that the published reports include all expected outcomes, including those that were pre-specified
Other bias	Unclear risk	Standard deviations not reported. No other source of potential bias detected

Bacchi-Modena 1997

Methods	Multicentre, double-blind, randomised, placebo-controlled trial Number of centres: 11 in Italy Number of women randomised: 109 Number of women analysed: 98 Statistical analysis: analysis of covariance (ANCOVA) with baseline value as covariate and treatment and centre as further effects in the model was used for the primary efficacy variable as well as for all secondary variables. For tests on treatment differences a significance level of 5% (two-sided) was applied Withdrawals and losses to follow-up: n = 11; 5 because of adverse events Duration of trial: 12 weeks Conflict of interest: not stated
Participants	Healthy female outpatients, requiring treatment for climacteric symptoms,were required to be at least 8 months after their last natural menstrual cycle or at least 6 weeks post-oophorectomy Inclusion Criteria: a minimum mean number of 7 moderate to severe hot flushes per 24 h during the last 2 weeks of a 4-week run-in period. Baseline serum FSH concentrations > 50 IU and serum E2 concentrations < 20 pg/mL Exclusion criteria: women who had received oral HT within 8 weeks, transdermal HT within 4 weeks or vaginal estrogen within 6 weeks before the start of trial treatment

Bacchi-Modena 1997 (*Continued*)

Interventions	Transdermal beta-estradiol 0.05 mg/d versus placebo
Outcomes	• Mean daily change from baseline of frequency and severity of moderate to severe hot flushes • Adverse experiences were assessed by direct questioning.
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not reported
Allocation concealment (selection bias)	Unclear risk	Not reported
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Authors only reported double blind study
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not reported
Incomplete outcome data (attrition bias) All outcomes	Low risk	Losses 11/109 (10%). ITT analysis
Selective reporting (reporting bias)	Low risk	The study protocol is not available but it is clear that the published reports include all expected outcomes, including those that were pre-specified
Other bias	Unclear risk	Standard deviations not reported. These were imputed by the review authors

Buster 2008

Methods	Multicentre, double-blind, randomised, placebo-controlled trial Number of centres: 43 in USA Number of women randomised: 458 Number of women analysed: 454 Statistical analyses were performed using SAS 8.2 (SAS Institute Inc., Cary, NC) on a PC platform. Statistical tests to evaluate treatment differences were two-tailed, with a significance level of 5% (0.05), and were declared statistically significant if the calculated P value was 0.05 or less. Treatment by region interaction analyses were performed at the 0.10 level. Adverse event terms were coded in Medical Dictionary for Regulatory Activities Version 7.1 in a one-way analysis of variance Duration of trial: 12 weeks Withdrawals and losses to follow-up: n = 81, 12 because of adverse events Conflict of interest: not stated
Participants	Participants were naturally or surgically postmenopausal women aged 35 years or older who were healthy, with an average of at least 8 moderate to severe hot flushes per day (56 or more per week) Inclusion Criteria: 12 months of spontaneous amenorrhoea or 6 months of amenorrhoea with serum FSH levels more than 40 IU/L or at least 6 weeks after, surgical bilateral oophorectomy with or without hysterectomy. Pap test with no dysplastic or malignant cells, screening mammogram and breast examination with no masses or other findings suspicious of malignancy, and endometrial biopsies showing no hyperplasia or cancer in women with a uterus Exclusion criteria: known hypersensitivity or known reaction to estrogens or progestins; clinically relevant disease that might preclude safe participation; use of progestin implants or injectable drug therapy or estrogen injectable or pellet therapy within 1 year, BMI more than 35 kg/m2; uncontrolled hypertension (diastolic blood pressure 95 mmHg or more and/or systolic blood pressure 180 mmHg or more); clinically relevant triglyceride levels 3.4 mmol/L or more (300 mg/dL or more); documented history of coagulopathy, thrombophlebitis, thrombosis, or thromboembolic disorders; history of cutaneous contact allergy to adhesives, cosmetics, or topical medications, including sunscreens; abnormal genital bleeding; and current dermatologic disease
Interventions	Beta-estradiol nasal spray 0.021 mg/d, 0.029 mg/d, and 0.040 mg/d for the one-, two-, and three-spray doses, respectively versus placebo
Outcomes	• Mean daily change from baseline of frequency and severity of moderate to severe hot flushes • Adverse events were reported by the participant
Notes	Study author reported only a washout period ".. from estrogen-containing medications before baseline assessments..."

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Study authors reported "The women were randomly assigned to treatment groups us-

Buster 2008 (*Continued*)

		ing a computer-generated central randomization schedule ...“
Allocation concealment (selection bias)	Low risk	Study authors reported ”...and the applicator number was assigned through an interactive voice response system...“
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Study authors reported ”All study site personnel and the participants were blinded as to the participant’s treatment group (active or placebo)...“, ”Applicators containing E2 or placebo were identical in appearance...“
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not reported
Incomplete outcome data (attrition bias) All outcomes	High risk	81/454 (17%) losses in the intervention group and unbalanced with control groups No ITT
Selective reporting (reporting bias)	Low risk	All expected outcomes and all outcomes listed in the protocol were reported
Other bias	Low risk	The study appears to be free of other sources of bias

Cohen 1999

Methods	Multicentre double-blind, randomised, placebo-controlled trial Number of centres: 9 in USA Number of women randomised: 259 Number of women analysed: 242 Statistical analysis: for primary outcomes the authors used an analysis of covariance model including treatment and centre as factors and centralised baseline as a covariate. Treatment-by-baseline and treatment-by-centre interaction terms were included in the model. Student t-tests were performed using an alpha = 0.05 level of significance to assess the difference in least significant means between intervention and control Withdrawals and losses to follow-up: 15 (6%) (5 receiving active treatment, 10 receiving placebo),14 participants (5 receiving active treatment, 9 receiving placebo) did not meet protocol criteria (participant’s diary recorded during the run-in period did not show 10 days of moderate-severe flushes), 1 patient receiving placebo used proscribed medication (oral beta-estradiol) Conflict of interest: none declared
Participants	Healthy, postmenopausal women aged 35 years or older Inclusion criteria: a minimum of 7 hot flushes per 24 hours or 60 hot flushes per week. Participants were required to have moderate to severe hot flushes during ≥ 10 days of the 2-week, drug-free run-in period (immediately before the initiation of study therapy)

	, as indicated in symptoms diaries. Postmenopausal status was confirmed on the basis of ≥ 12 months' amenorrhoea or a history of bilateral oophorectomy regardless of serum estradiol (E2) or FSH levels. Participants with 6-11 months of amenorrhoea and those who had been hysterectomised before cessation of menses were required to have serum E2 values ≤ 20 pg/mL and FSH values ≥ 40 mIU/mL A prestudy washout of 4 weeks for transdermal administration, 8 weeks for oral administration, and 6 months for injection or implants was required for participants receiving previous HRT Exclusion criteria: clinically significant findings involving the reproductive system (e.g. undiagnosed vaginal bleeding, abnormal Papanicolaou smear, endometrial thickness > 5 mm, or polyps) or any malignancy, uncontrolled thyroid disease, cardiovascular or other medical complications that might confound study assessments
Interventions	0.0375 mg/d beta-estradiol transdermal patch (11 cm2) or a matching placebo
Outcomes	• Mean daily change from baseline of frequency and severity of moderate to severe hot flushes • Adverse events were recorded in participant diaries at each study visit
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not reported
Allocation concealment (selection bias)	Unclear risk	Not reported
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Authors only reported double-blind study. No further details
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not reported
Incomplete outcome data (attrition bias) All outcomes	Low risk	15/257 (6%), balanced dropouts in all group. ITT analysis
Selective reporting (reporting bias)	Low risk	The study protocol is not available but it is clear that the published report includes all expected outcomes, including those that were pre-specified
Other bias	Low risk	The study appears to be free of other sources of bias

De Aloysio 2000

Methods	Single centre, double-blind, randomised, placebo-controlled trial Number of centres: 1, Bologna, Italy Number of women randomised: 156 Number of women analysed: not stated Statistical analysis: the comparability between groups at the end of the 2-week baseline for the mean number of hot flushes per day was tested by ANOVA, with Tukey's HSD test for multiple comparisons. The statistics of the Kupperman Index were based on the last observation carried forward (LOCF). The responder rate at the end of the treatment was analysed by the Chi^2 test on the ITT participants The secondary end points of efficacy were also tested by ANOVA, with Tukey's HSD test for multiple comparisons. The Kruskal-Wallis test was used for the participant's efficacy judgements. All comparisons between the 3 groups were performed considering a 5 % α significance level. The Bonferroni approach to multiple comparisons was implemented Withdrawals and losses to follow-up: n = 20.3 because of adverse events Duration of trial: 12 weeks Conflict of interest: not stated
Participants	Inclusion criteria: women attending the centre because of climacteric symptoms, in natural or surgical menopause, aged over 35 years with at least 6 months since the last regular menstruation in case of natural menopause, or at least 4 weeks from ovariectomy, with serum FSH $\geq$ 40 mIU/mL and/or E2 < 30 pg/mL, and with an average number of at least 5 hot flushes per day Exclusion criteria: presence or history of possibly estrogen dependent tumours, undiagnosed vaginal bleeding, history of endometriosis, history of thrombophlebitis, thrombosis or thromboembolic disease, hepatic or renal failure, severe and uncontrolled concomitant diseases, presence or history of skin diseases or allergies that could affect the local tolerability of the patch or the absorption of estradiol and HT by injection or implants during the previous 12 months
Interventions	Transdermal beta estradiol 0.025 or 0.0375 mg/d versus placebo
Outcomes	• Mean daily change from baseline of frequency and severity of moderate to severe hot flushes • a) systemic adverse events including ○ type, severity, onset and duration and possible causal relationship with the medication; ○ presence and severity of breast tenderness; ○ possible episodes of vaginal bleeding; • local tolerability was assessed at each clinic visit; • possible adverse reaction of the skin at the site of patch application was graded from 0 (absence of reaction) to 9 (severe skin lesions)
Notes	13 patients (25 %) in the placebo group did not complete the study Data on effectiveness are skewed, reported in additional tables

Risk of bias

Bias	Authors' judgement	Support for judgement

De Aloysio 2000 (*Continued*)

Random sequence generation (selection bias)	Low risk	Authors reported that "...eligible women were enrolled and blindly assigned according to a computer generated randomization list to the following 3 parallel treatment groups..."
Allocation concealment (selection bias)	Low risk	Authors reported that "..eligible women were enrolled and blindly assigned according to a computer generated randomization list to the following 3 parallel treatment groups.."
Blinding of participants and personnel (performance bias) All outcomes	Low risk	The E2 patches were manufactured in a similar way, were without enhancers or other excipients, had the same composition of adhesive and were put in identical sachets to assure blindness. Placebo: transdermal patch identical to the previous ones but without E2"
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not reported
Incomplete outcome data (attrition bias) All outcomes	High risk	Losses: 20/156 (13%). High rate of losses in the placebo group (33%) compared to the intervention group (9.6% and 4%). No ITT analysis
Selective reporting (reporting bias)	Low risk	The study protocol is not available. Includes all expected outcomes, including those that were pre-specified
Other bias	Low risk	The study appears to be free of other sources of bias

De Vrijer 2000

Methods	Multicentre, double-blind, randomised, placebo-controlled trial Number of centres: 16 in The Netherlands Number of women randomised: 254 Number of women analysed: 227 Statistical analysis: the primary efficacy criterion was analysed by analysis of covariance (ANCOVA), with baseline values as covariate, and treatment and centre as factors. Exploratory analyses of secondary efficacy criteria (daytime hot flushing and Kupperman Index) were also assessed by ANCOVA. Night sweats were summarised descriptively. All statistical tests were carried out at the 2-sided 5% level Withdrawals and losses to follow-up: 6 discontinued due to poor efficacy, 18 due to adverse experiences and 3 due to administrative problems Duration of trial: 12 weeks Conflict of interest: not stated
Participants	Healthy postmenopausal female outpatients, with or without previous HT, requiring treatment for climacteric symptoms. Women with at least 8 months after last spontaneous menstrual bleeding or 6 weeks post-oophorectomy, with a mean number of seven or more moderate to severe hot flushes (including night sweats) per 24 h in the 14 days prior to randomisation. In addition, all participants had to have plasma FSH and estradiol levels in the normal postmenopausal range (i.e. > 50 IU/L and < 20 pg/mL, respectively) Exclusion criteria: women with contraindications for HT, with recent HRT (oral, transdermal or vaginal) in previous 8 weeks, implant in preceding 12 months), or with any previous unopposed estrogen use for more than 3 months
Interventions	Transdermal beta-estradiol 0.05 or 0.1 mg/d versus placebo
Outcomes	• Mean daily change from baseline of frequency and severity of moderate to severe hot flushes • Frequency and severity of any adverse event
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Author reported "...patients who met the inclusion criteria were randomized in equal numbers to receive either E2 Matrix 0.10 mg, E2 Matrix 0.05 mg, or placebo..."
Allocation concealment (selection bias)	Unclear risk	No specific discussion of allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Author reported "As the two-dosing systems were different in size in order to maintain blindness a double-dummy design was used..."

De Vrijer 2000 (*Continued*)

Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not reported
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	27/254 (11%) women withdrew or lost to follow-up. ITT analysis
Selective reporting (reporting bias)	Low risk	No protocol available but study reports all expected outcomes
Other bias	Low risk	The study appears to be free of other sources of bias

Good 1999

Methods	Multicentre, double-blind, randomised, placebo-controlled trial Number of centres: 26 USA Number of women randomised: 321 Number of women analysed: 321 Statistical analysis: all statistical tests were conducted as two-tailed with a critical probability of 0.05 for declaring significance. Pearson's Chi-squared test or Fisher's exact probability test were used to examine the homogeneity of study groups with respect to nominal variables. For continuous variables, analysis was performed with study drug as a factor. Where analysis of variance (ANOVA) assumptions were not met, the Kruskal-Wallis non-parametric test was used to compare study-drug groups Withdrawals and losses to follow-up: n = 47, 16 because of adverse events Duration of trial: 12 weeks Conflict of interest: reported source of funding per TheraTech Inc
Participants	Healthy postmenopausal women Inclusion criteria: age at least 21 years if surgically menopausal, 45 years if naturally menopausal; amenorrhoeic for at least 6 months; experiencing 60 or more moderate to severe hot flushes per week; serum estradiol concentrations $\leq$ 73 pmol/L and FSH concentrations $\geq$ 40 mIU/mL Exclusion criteria: previous serious dermatological disease or active skin disorder; any psychiatric or psychological illness at time of entry; clinically significant abnormalities in medical history, physical examination or mammography; or the use of any investigational drug within the previous 60 days. Prospective participants were also excluded on the basis of the following medical data obtained at the first 2 screening visits: elevated sitting blood pressure (systolic > 165 mmHg or diastolic > 95 mmHg); any degree of dyskaryosis in a cervical Pap smear; any significant abnormalities in clinical laboratory tests; or a body weight > 130% of the participant's ideal range according to the 1983 Metropolitan Height and Weight Table for Women
Interventions	Transdermal beta-estradiol 0.05 or 0.1 mg/d versus (CEE) Conjugated Equine estrogen 0.625mg/d or 1.25mg/d

Good 1999 (*Continued*)

Outcomes	• Mean weekly change in frequency and severity of moderate to severe vasomotor symptoms from baseline • Monitoring adverse events
Notes	Period of washout: abstention from previous HT for at least 6 weeks prior to the first screening visit (12 months in the case of estrogen implants)

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not reported
Allocation concealment (selection bias)	Unclear risk	Not reported
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Author reported double-dummy "Administration of oral CEE was blinded by encapsulation of commercially acquired tablets (Premarin 0.625 mg) as either one or two tablets per capsule. Each patient received simultaneously an 1 8-cmz transdermal system and a 36-cm2 transdermal system twice weekly and swallowed a capsule daily. .."
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	No specific discussion of blinding of outcome assessment
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Losses 47/321 (14.6%) No ITT analysis
Selective reporting (reporting bias)	Low risk	No protocol available but the study reports all expected outcomes
Other bias	Unclear risk	Findings on effectiveness reported only as a graph, so unsuitable for analysis. Reported source of funding per TheraTech Inc

Gordon 1995

Methods	Two multicentre studies, double-blind, randomised, placebo-controlled trial Number of centres study 1: 25 centres in USA Number of centres study 2: 34 centres in USA Number of women randomised study 1: 214 Number of women analysed study 1: 191 Number of women randomised study 2: 390 Number of women analysed study 2: 390 Statistical analysis: baseline weekly mean hot-flushes rates were compared between treatments with an analysis of variance F test and a Wilconxon rank sum test. Analysis of variance was performed on the change in weekly hot-flush rate and subjects and investigators' global assessments for all participants included in the efficacy analysis. Tukey's multiple comparison test was used to test significant differences between treatments. A P value of < 0.05 was considered significant Withdrawals and losses to follow-up study 1: n = 50. 22 because of adverse events Withdrawals and losses to follow-up study 2: n = 64. 32 because of adverse events (15 in the 0.05 mg/d, 10 in the mg/d, and 7 in the CEE groups) Duration of trial: 11 weeks Conflict of interest: not stated
Participants	Postmenopausal women. Natural menopause participants had not menstruated for at least 12 months and surgical menopause participants had undergone bilateral oophorectomy at least 4 weeks prior to inclusion Inclusion criteria: only women with no clinically significant medical history and no contraindication to HT were considered for inclusion, normal Papanicolau smear, normal endometrial biopsy (if she had not had a hysterectomy), baseline serum estradiol level less than 20 pg/ mL, and baseline serum FSH level greater than 50 mIU/mL. In order to enter the 11-week treatment phase of the study, potential subjects must have experienced a minimum of 5 moderate to severe hot flushes per week, or a minimum of 15 hot flushes of any severity per week, for 2 consecutive weeks Exclusion criteria: women with no clinically significant medical history and contraindication for HT. Subjects qualified after having performed a examination physical, an electrocardiogram and laboratory tests (complete blood count, serum electrolytes, liver and renal function tests, total protein and albumin, cholesterol, pregnancy test if appropriate and urinalysis) Washout: Injectable HT had to be discontinued 12 weeks prior to entering the qualification period and progesterone and other steroid compounds capable of interacting with the study medications were prohibited during these trials
Interventions	Transdermal beta-estradiol 0.05 or Transdermal beta-estradiol 0.1 versus placebo (study 1) Transdermal beta-estradiol 0.05 versus 0.1 mg/d versus (CEE) conjugated Equine estrogen 0.625 mg/d (study 2)
Outcomes	• Mean weekly change in frequency and severity of moderate to severe vasomotor symptoms from baseline • Any reactions on site of application, lessening of patch adhesion
Notes	Highest withdrawal rates were in placebo (30%) and 0.05 mg beta-estradiol groups (26%) vs 0.1 mg beta-estradiol group (13%) because of inadequate therapeutic response (P value < 0.05) (study1)

Gordon 1995 (*Continued*)

	Data unsuitable for analysis reported in an additional table	
Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not reported
Allocation concealment (selection bias)	Unclear risk	Authors reported "At each centre , qualified subjects in each study were randomly assigned to one of three treatments groups'
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Authors reported "To preserve blinding subjects randomized to the 0.05 mg estradiol group also wore 25 cm placebo patch. .." (study 1) . "Blinding was preserved in a manner similar to that employed in study 1..." (study 2)
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not reported
Incomplete outcome data (attrition bias) All outcomes	High risk	Losses to follow-up: 50/191 study 1 (26% (30% losses in the intervention group and unbalanced with control groups.)) 64/390 study 2 (16%) No ITT analysis
Selective reporting (reporting bias)	Unclear risk	No protocol available, limited reporting of adverse events
Other bias	Low risk	No other source of potential bias detected

Haines 2009

Methods	Multicentre, double-blind, randomised, placebo-controlled trial Number of centres: 5 centres in Asia Number of women randomised: 165 Number of women analysed: 161 Statistical analysis: the relative change in the frequency of hot flushes from baseline to week 12 was compared between treatment groups using a 2-sided Wilcoxon rank-sum (Mann-Whitney) test. The null hypothesis tested was the assumption of equal distributions of the primary target variable in both treatment groups Withdrawals and losses to follow-up: n = 14, 4 in the E2 group and 10 in the placebo group, because 5 never started treatment or had unknown status, 1 adverse event, 3 for withdrawal of consent, 4 were lost to follow-up and one discontinued for 'other' reasons Duration of trial: 12 weeks Conflict of interest: not stated
Participants	Women were eligible for the study if they were aged between 40 and 65 years, had undergone natural menopause Inclusion criteria: postmenopausal status was confirmed on the basis of 12 months' amenorrhoea or 6 months' amenorrhoea with serum FSH $\geq$ 40 m IU/mL or bilateral oophorectomy (6 weeks postsurgery) and had at least 24 hot flushes (of any severity) within a 7-day screening period Exclusion criteria: recently used estrogen-containing products, an abnormal cervical smear test, endometrial thickness of 5.0 mm, any condition that could interfere with study medication or interpretation of results, concomitant use of inducers or inhibitors of CYP3A4 or drugs effective in treating hot flushes, received anticoagulant treatment for the past 6 months, or known severe dyslipoproteinaemia
Interventions	Transdermal patch 0.014 mg beta-estradiol/day versus placebo
Outcomes	• Mean weekly change in frequency of moderate to severe vasomotor symptoms from baseline • Adverse events; clinical laboratory evaluation • Quality of life
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Authors reported "Randomization was by a centrally provided computer-generated list. .."
Allocation concealment (selection bias)	Unclear risk	Not reported
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Authors only reported double-blind study, no further details

Haines 2009 (*Continued*)

Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not reported
Incomplete outcome data (attrition bias) All outcomes	Low risk	Losses: 14/165 (8%)
Selective reporting (reporting bias)	Low risk	No protocol available, but all expected outcomes reported
Other bias	Unclear risk	Findings on effectiveness reported only as a graph, so unsuitable for analysis

Honjo 2009

Methods	Multicentre study, double-blind, randomised, placebo-controlled, three-parallel-group trial Number of centre: 23 centres in Japan Number of women randomised: 211 Number of women analysed: 211 Statistical analysis: descriptive statistics were used to assess the data. Reductions in the number of hot flushes with each treatment were compared using a 2-tailed t-test and Holm's step-down method; secondary endpoints were compared using the Wilcoxon test, Fisher's exact test and Holm's step-down method. Significance was set at 5%. Demographic and baseline data were compared using the w2 test for categorical data and one-way analysis of variance (ANOVA) for continuous data; significance was set at 15% Withdrawals and losses to follow-up: n = 7, 4 in the placebo, 2 in the E2 0.5 mg and one in the E2 1.0 mg group. No subjects discontinued due to adverse events Duration of trial: 8 weeks Conflict of interest: none reported
Participants	Women with naturally occurring climacteric symptoms or ovarian deficiency symptoms following bilateral oophorectomy Inclusion criteria: women were eligible for the study if they were aged at least 40 but less than 65 years, had experienced either natural menopause (defined as 1 year since last menses or serum E2 $\leq$ 20 pg/mL and FSH $\geq$ 30 mIU/ mL) or bilateral oophorectomy (1 month postsurgery) and had climacteric or ovarian deficiency symptoms, defined as an average of at least 3 moderate or severe hot flushes per day in the 7 days before the start of study drug administration Exclusion criteria: women were excluded for gynaecologic (including cancer) or psychological disorders, hypertension, insulin-treated diabetes, migraine or epilepsy, history of thromboembolism or cardiovascular disease, or recent treatment with sex steroid hormones or autonomic nervous system regulators
Interventions	Micronised beta-estradiol/day 0.5 or 1.0 mg versus placebo tablets once daily

Outcomes	• Mean daily change from baseline of frequency and severity of moderate to severe hot flushes • Adverse events (including genital bleeding)
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Authors reported ".. Randomization was by six subject batch in each centre, with each batch including two subjects in each treatment group; consecutive subjects were allocated to treatment by random number within the batch..."
Allocation concealment (selection bias)	Unclear risk	No specific discussion on allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Authors reported **"...**double-blind..." No further details
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	No specific discussion on blinding of outcome assessment
Incomplete outcome data (attrition bias) All outcomes	Low risk	Losses: 7/211 (3.3%)
Selective reporting (reporting bias)	Low risk	No protocol available, but reports all expected outcomes
Other bias	Low risk	The study appears to be free of other sources of bias

Nielsen 2006

Methods	Multicentre double-blind, randomised, placebo-controlled trial Number of centres: 2 in Denmark Number of women randomised: 335 Number of women analysed: 268 Treatment effect was studied on an ITT basis using the change in score from W0 to the last questionnaire available i.e. up to W26 for vasomotor symptoms and up to W104 for all other dimensions. A global score was established from linear transformation of the corresponding answers. Between-group differences in mean change scores were evaluated with a non-parametric covariance analysis. Two-sided 95% CI of the differences were calculated Duration of trial: 2 years Withdrawals and losses to follow-up: 67/335 (20%) Conflict of interest: none declared
Participants	Women, 40-65 years of age at baseline,who were less than 5 years past menopause Ethnicity: white Inclusion criteria: amenorrhoea for more than 12 months or for more than 6 months and a concomitant serum level of estradiol below 0.16 nmol/L and FSH level above 42 IU/L. All women who had undergone hysterectomy had menopause confirmed by determination of serum estradiol and FSH at least 2 months prior to study Surgical menopause (bilateral ovariectomy): operation performed at least 6 weeks before study Women had to be osteopenic (BMD T score < −1) and not to complain of severe climacteric symptoms
Interventions	S21400 (intranasal beta-estradiol) micronised 150 µg per day, S21400 300 µg per day +, or placebo or S21400 (intranasal beta-estradiol) micronised 150 µg per day, S21400 300 µg per day +, progesterone 200 mg/d, 14 days out of 28, or placebo (women with an intact uterus)
Outcomes	• QoL was assessed based on the validated Women's Health Questionnaire designed for peri- and post-menopausal women. Change from baseline scores • Any adverse events reported by the participant during the study
Notes	Methods described in another publication (Nielsen 2004) Data skewed, reported in an additional table

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	The treatment allocation lists were drawn up and encoded by the Institut de Recherches Internationales Servier
Allocation concealment (selection bias)	Unclear risk	The treatment allocation lists were drawn up and encoded by the Institut de Recherches Internationales Servier

Nielsen 2006 (*Continued*)

Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Authors only reported double-blind study. No further details
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Double-blind. No further details
Incomplete outcome data (attrition bias) All outcomes	High risk	Losses: 67/335 (20%). ITT analysis
Selective reporting (reporting bias)	Low risk	No protocol available, but reports all expected outcomes
Other bias	Low risk	The study appears to be free of other sources of bias

Notelovitz 2000a

Methods	Study design: parallel, double-blind, multicentre, placebo-controlled, randomised clinical trial Randomisation: block randomisation within each centre Number of centres: 15 in the USA Duration of trial: 3 months Power calculations: yes (based on moderate to severe hot flush outcome) Number of women randomised: 333 in to 1 of 5 treatment groups Number of women analysed: 280 for moderate/severe hot flushes and 324 for hot flush weekly weighted score ITT analysis: no for moderate/severe hot flushes, yes for hot flush weekly weighted score Losses to follow-up/withdrawals from treatment: 53/333 = 16% for moderate/severe hot flushes and 4/333 = 1.2% for hot flush weekly weighted score Compliance: not stated Source of funding: supported by grants to participating institutions and Novo Nordisk Pharmaceuticals Inc
Participants	Menopausal status: peri- and post-menopausal Age: mean 54.12 ± 4.14 years (mean ± SD) (range 40-60 years) Ethnicity: race (n, %) in placebo, 0.25 mg E2, 0.5 mg E2, 1.0 mg E2, 2.0 mg E2 respectively: White - 60 (91%), 62 (91%), 61 (95%), 57 (85%), 62 (91%) Black - 3 (5%), 3 (4%), 2 (3%), 5 (7%), 1 (1%) Hispanic - 1 (2%), 2 (3%), 1 (2%), 4 (6%), 3 (4%) Asian/Pacific - 2 (3%), 0 (0%), 0 (0%), 1 (1%), 2 (3%) Other - 0 (0%), 1 (1%), 0 (0%), 0 (0%), 0 (0%) Source: study population was obtained from the investigators´ sites or through local advertising (i.e. a mixture of clinical and general population) Inclusion criteria: menopause symptoms persisting for more than 6 months, healthy menopausal women with an intact uterus, 40-60 years old, at least 56 moderate-severe

	hot flushes/week, at least 6 months amenorrhoea, E2 levels ≤ 20 pg/mL, FSH > 50 IU/L Exclusion criteria: history of endometrial hyperplasia, abnormal bleeding of unknown origin, endometrial thickness at least 5 mm, history of estrogen-dependent tumours, gallbladder, liver kidney or endocrine diseases except controlled thyroid disease, venous thromboembolism, cerebrovascular accidents, myocardial infarction or ischaemic heart disease, history of severe headache or migraines, high blood pressure, alcohol or drug abuse, smoking > 15 cigarettes/day, weight increased more than 20% over ideal body weight, use of steroid hormones/drugs known to influence estrogen metabolism & use of HRT within 2 months prior to randomisation Confirmation of ovarian failure: at least 6 months amenorrhoea, E2 levels ≤ 20 pg/mL, FSH ≥ 50 IU/L Baseline equality: matched for age, time of amenorrhoea, weight, baseline hot flush symptoms Baseline symptoms: At least 56, with 72 ± 21 (mean ± SD) moderate-severe hot flushes/week & mean hot flush weekly weighted score 183 ± 61 (mean ± SD)
Interventions	Micronized beta-estradiol 0.25 or micronized beta-estradiol 0.5 mg/d or beta-estradiol 1.0 mg/d or micronized beta-estradiol 2.0 mg/d versus placebo. The HT and placebo preparations were identical in appearance Co-interventions: none reported
Outcomes	• Mean weekly change in frequency and severity of moderate to severe vasomotor symptoms from baseline • Adverse event frequency
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Authors reported that "The randomization code was generated centrally using a block size of five to ensure equal distribution of treatment groups."
Allocation concealment (selection bias)	Unclear risk	Authors reported that "On entry, subjects were assigned to the lowest available randomization number for each site."
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Authors reported that "All study drugs were identical in appearance and packaging, and were manufactured and supplied by Novo Nordisk A/S, Bagsvaerd, Denmark."
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Author reported that "...serum samples collected during therapy were analyzed at the end of the study to preserve masking of

Notelovitz 2000a (*Continued*)

		treatment allocation of all centres and personnel related to conduct of this study."
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Losses 53/333 (16%)
Selective reporting (reporting bias)	Low risk	No protocol available, but reports all expected outcomes
Other bias	Unclear risk	Findings on effectiveness reported only as a graph, so unsuitable for analysis. Supported by grants to participating institutions and Novo Nordisk Pharmaceuticals Inc

Notelovitz 2000b

Methods	Multicentre, double-blind, randomised, placebo-controlled trial Number of centres: 23 in USA Number of women randomised: 145 Number of women analysed: 145 Statistical analysis: all statistical testes were two-tailed and conducted at the 5% significance level. The percentage change in number of hot flushes from baseline to the end for treatment (or last available measurement) was calculated for each treatment group and was compared with the placebo group using a 2-way fixed-effect analysis of variance. This was also used to compare the percentage change from baseline to the end of treatment in the mean number of moderate or severe hot flushes per day. Comparisons between active treatment groups versus placebo for subjects free of hot flushes were made using Fisher's exact test. For mean daily hot flushes per week analysis of covariance was used; P values were calculated from t-tests on least square means Withdrawals and losses to follow-up: n = 23. Withdrawals because of adverse events not reported Duration of trial: 12 weeks Conflict of interest: none declared
Participants	Healthy postmenopausal women who exhibited moderate to severe vasomotor symptoms. The following groups were included: naturally postmenopausal women aged 40-60 years (women who had not experienced menses for at least 12 months before start of the study), women who had undergone hysterectomy, and women aged 25-60 years who had undergone bilateral oophorectomy with or without hysterectomy (eligible 3 weeks after surgery) Inclusion criteria: presence of an average of 8 or more moderate to severe hot flushes per day during the 14 consecutive days before the start of drug administration, FSH hormone levels were required to be $\geq$ 40 mIU/L and $\geq$ 30 mIU/L for subjects with and without prior HT, respectively, whose menopause was natural or surgically induced, serum estradiol levels had to be $\leq$ 20 pg/ml Exclusion criteria : exposure to injectable or implantable sex steroids within 6 months of admission and hypersensitivity to study drug ingredients
Interventions	Oral beta-estradiol 0.5 or 1 mg/d versus placebo

Notelovitz 2000b (*Continued*)

Outcomes	• Mean daily change from baseline of frequency and severity of moderate to severe hot flushes • Adverse events were recorded by the investigator at each visit
Notes	Period washout: abstention from previous HT for at least 6 weeks before the pre-randomisation visit and 8 weeks before the randomisation visit

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Authors reported "randomized"
Allocation concealment (selection bias)	Unclear risk	Not reported
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Authors reported "Double-blind". No further details
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	No specific discussion on blinding of outcome assessment
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Losses 23/145 (16%) ITT analysis
Selective reporting (reporting bias)	Low risk	Includes all expected outcome
Other bias	Unclear risk	Authors reported "The protocol was amended because of unfavorable rates of bleeding/spotting noted with combinations of 2 mg E_2 and NGM used in comitant studies."

Rovati 2000

Methods	Multicentre study with a prospective, randomized, double-blind, placebo-controlled, parallel-group Number of centres: 15 in Italy Number of women randomised: 311 Number of women analysed: 277 Statistical analysis: Frequency of patient dropouts and/or use of rescue medication were compared using the x2 test or Fisher's exact test, as appropriate. Groups were compared at the end of the 2-week baseline for the primary efficacy outcome, i.e. the mean number of hot flushes per day, using single-factor ANOVA, with Tukey's HSD test for multiple comparisons Withdrawals and losses to follow-up: n = 34, 5 because of adverse events Duration of trial: 12 weeks Conflict of interest: not stated
Participants	Women in natural or surgical menopause Inclusion criteria: women with at least 6 months since the last regular menstruation in the case of natural menopause or at least 8 weeks from ovariectomy, with serum FSH > 50 mIU/mL and E2 < 20 pg/mL, and with an average of at least seven moderate to severe hot flushes per day Exclusion criteria: participants excluded were those with known or suspected breast cancer; history of possible estrogen-dependent tumours; pathological results from the cervical smear performed at enrolment, any mammography examination, endometrial biopsy or vaginal ultrasound; undiagnosed genital bleeding; history of endometriosis; active thrombophlebitis or thromboembolic disorders; hepatic or renal failure; severe and uncontrolled concomitant diseases; presence or history of skin diseases or allergies that could affect the local tolerability of the patch or the absorption beta-estradiol ,alcohol or drug abuse; psychiatric disorders; or any other factor limiting the ability of participants to understand the aim/procedures or to give their informed consent
Interventions	Transdermal beta-estradiol 0.025 mg/d or 0.050 mg/d versus placebo
Outcomes	• Mean daily change from baseline of frequency and severity of moderate to severe hot flushes • Systemic adverse events with their type, severity, onset and duration, and possible causal relationship with medication
Notes	Washout period: HT by injection or implant during the previous 12 months; women with a uterus who had received unopposed estrogen treatment for more than 3 months Data on effectiveness skewed and unsuitable for analysis. Reported in an additional table

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Authors reported "...eligible women were enrolled and blindly assigned according to a computer-generated randomization list to the following four parallel treatment groups..."

Rovati 2000 (*Continued*)

Allocation concealment (selection bias)	Unclear risk	No specific discussion on allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Authors reported "The seven-day patches were manufactured in a similar way, are without absorption enhancers or other excipients, have the same matrix composition and were put in identical sachets to assure double blindness..." In case of poor symptom control, Derm-50 was used as open-label rescue medication" Potential introduced by use of open-label rescue medication; there was between-arms difference in the number of women who opted for rescue
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	As above
Incomplete outcome data (attrition bias) All outcomes	High risk	Losses 34/277 (11%) 21% of the placebo group dropped out to move to open-label medication (for this reason) vs 8% in the 0.025 mg/d or 3% 0.050 mg/d transdermal beta-estradiol groups. ITT analysis
Selective reporting (reporting bias)	Low risk	No protocol available, but reports all expected outcomes
Other bias	Low risk	No other source of potential bias identified

Simon 2006

Methods	Multicentre, randomised, double-blind, placebo-controlled trial Number of centres: 37 in USA Number of women randomised: 200 Number of women analysed: 200 Statistical analysis: Between-group differences in mean reduction from baseline in frequency of hot flush counts were assessed using analysis of covariance (ANCOVA). Treatment efficacy was assessed on the ITT population, defined as all randomised participants who received one or more doses of study medication. Tests comparing treatment groups were 2 sided and performed at the 0.05 level of significance Withdrawals and losses to follow-up: n = 17 (10 participants receiving MNPEE and 7 receiving placebo emulsion) 5 because of adverse events Duration of trial: 12 weeks Conflict of interest: not declared

Simon 2006 (*Continued*)

Participants	Postmenopausal women with 7 or more moderate to severe hot flushes per day Inclusion criteria: more than 12 months of amenorrhoea or more than 6 months of amenorrhoea if accompanied by serum levels of FSH > 40 mIU/mL and serum estradiol < 20 pg/mL. Women who had experienced 120 or more moderate to severe hot flushes over 2 weeks of the screening period, hysterectomy, either with or without bilateral oophorectomy, were eligible for the study if surgery was performed 1 year or more before study entry, normal or benign pelvic examination and Papanicolaou smear, and a normal or benign mammogram within 9 months before study entry Exclusion criteria: participants were ineligible if they had received oral or transdermal HT within 2 months before study entry, within 3 months for implantable HT, and within 6 months for injectable HT. Women with a history of breast, cervical, or uterine cancer or any other malignancy and chronic dermatologic conditions that required topical maintenance therapy, a history of cerebral vascular disease, coronary artery disease, myocardial infarction, uncontrolled hypertension, undiagnosed abnormal vaginal bleeding, endocrine disease other than controlled thyroid disease on stable doses of therapy, and allergy or hypersensitivity either to estrogens or to any constituents of this topical medication
Interventions	Micellar nanoparticle beta-estradiol emulsion (MNPEE containing 8.6mg of beta-estradiol) versus placebo
Outcomes	• Mean daily change from baseline of frequency and severity of moderate to severe hot flushes • Adverse events were assessed throughout the study and were categorised using standard criteria
Notes	Data on effectiveness skewed. Reported in an additional table

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Authors reported "...were then randomized to receive either MNPEE or placebo for 12 weeks."
Allocation concealment (selection bias)	Unclear risk	No specific discussion on allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Authors reported "Micellar nanoparticle estradiol emulsion (containing 8.6 mg of 17B -estradiol in 3.45 g of emulsion) or placebo was administered in identical foil-packed pouches..."
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	No specific discussion on blinding of outcome assessment

Simon 2006 (*Continued*)

Incomplete outcome data (attrition bias) All outcomes	Low risk	Losses: 17/200 (8.5%), ITT analysis
Selective reporting (reporting bias)	Low risk	No protocol available but reports all expected outcomes
Other bias	Low risk	No other source of potential bias detected

Simon 2007

Methods	Multicentre, randomised, double-blind, placebo controlled, parallel-group trial Number of centres: 28 sites in USA and 2 sites in Canada Number of women randomised: 484 Number of women analysed: 457 Statistical analysis: for all efficacy variables an adjustment for multiple testing was performed using Dunnett's test. Comparison of the treatment groups with respect to other quantitative variables were based on a 2-way analysis of variance model, including factors for centre, treatment, and centre-by-treatment interaction. For either model, the interaction term was removed from the final model if not statistically significant (P > 0.10). Each dose group was compared with placebo using Dunnett's test, with comparisons based on the least squares means derived from the analysis of covariance or analysis of variance. Comparability of treatment groups with respect to categorical variables was based on the Cochrane Mantel-Haenszel general association statistic with centre as the stratification factor followed by the Mantel-Haenszel test if a statistically significant treatment group difference was found Withdrawals and losses to follow-up: n = 27. 9 because of adverse events Duration of trial: 12 weeks Conflict of interest: none declared
Participants	Women aged 18 years or older who had undergone natural (amenorrhoea for 12 months or more) or surgical menopause (bilateral oophorectomy with or without hysterectomy) 6 months or more before screening Inclusion criteria: serum estrogens less than 20 pg/mL, FSH more than 40 IU/mL, and BMI of 18-35 kg/m2. Eligible participants recorded the number, time, and severity (0. none; 1, mild; 2, moderate; or 3, severe) of hot flushes over the first 14 days of a 3-4-week screening period. Participants who experienced 60 or more moderate to severe hot flushes each week and were otherwise eligible had blood taken to measure serum E2 and entered a 1-week single-blind placebo lead-in period, during which they applied placebo gel once daily Exclusion criteria: women with history of estrogen-dependent neoplasia; endometrial hyperplasia; active hepatic, gallbladder, renal, or endocrine disease other than controlled thyroid abnormalities; or if they were receiving concomitant medications that could potentially interfere with hot flush frequency, severity, or their assessment
Interventions	Beta-estradiol gel 0.87 g/d (= 0.52 mg/d estradiol) or beta estradiol gel 1.7 g/d (= 1.02 mg/d estradiol) or beta-estradiol gel 2.6 g/d (= 1.56 mg/d estradiol/) versus placebo

Outcomes	• Mean daily change from baseline of frequency and severity of moderate to severe hot flushes • The overall incidence of treatment-emergent adverse events were reported
Notes	Data on effectiveness skewed; reported in an additional table

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Authors reported that participants "... were randomly assigned to groups and entered the 12-week double blind treatment period..." "Initially, participant numbers were randomly allocated in blocks of three and in a 1:1:1 ratio to E2 gel 1.7 g/d (1.02 mg E2), E2 gel 2.6 g/d (1.56 mg E2) or placebo gel using a computer-generated randomization list..."
Allocation concealment (selection bias)	Unclear risk	Study co-ordinators assigned treatment numbers to participants entering the 12-week double-blind treatment period
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Authors reported that "Participants, investigators, and outcome assessors were blinded to treatment assignment, and no unblinding occurred during the trial... participants applied the study drug from two metered-dose bottles (0.87 g gel per pump actuation), one labelled Bottle A and the other Bottle B, that contained placebo gel or E2 gel as necessary to deliver the appropriate treatment and dose."
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Authors reported that "Participants, investigators, and outcome assessors were blinded to treatment assignment, and no unblinding occurred during the trial..."
Incomplete outcome data (attrition bias) All outcomes	Low risk	Losses: 27/484 (5.6%) 7% discontinued in the placebo group, 3% discontinued in the 0.87 g group, 7% discontinued in the 1.7 g group. ITT analysis
Selective reporting (reporting bias)	Low risk	No protocol available but reports all expected outcomes

Simon 2007 (*Continued*)

Other bias	Low risk	No other source of potential bias detected

Speroff 1996a

Methods	The trial publication reports two independent RCTs. Speroff 1996a is reported as Study 1 in the publication Multicentre randomised, double-blind, placebo-controlled trial Number of centres: 17 in USA Number of women randomised: 213 Number of women analysed: 170 Statistical analysis: analysis of covariance was used to compare the frequency of hot flushes per week between treatment groups within each study for each of the 12 weeks of the study. Estradiol-estrone ratios of mean observed concentrations were calculated at each sampling time Duration of trial: 12 weeks Withdrawals and losses to follow-up: 43/213 (20%). High rate of losses in the placebo group (30%) compared to the intervention group (7% and 11%) Conflict of interest: this study was supported by Warner Chilcott, a division of Galen Holdings PLC, which has developed this product
Participants	Naturally menopausal women, all with prior hysterectomy Inclusion criteria: women with 56-140 hot flushes per week, had undergone hysterectomy, were at least 50 years old and naturally menopausal or at least 35 years old and surgically menopausal (bilateral oophorectomy at least 1 month before entry), had serum E2 concentrations of 20 pg/mL or less, and had FSH levels of 50 mlU/mL or more Exclusion criteria: women with contraindications to beta-estradiol replacement therapy or with a skin condition that might be exacerbated by the use of a transdermal system or that might mask a dermatologic reaction to the system
Interventions	Transdermal system delivering 0.02 mg and 0.04 mg of beta-estradiol versus placebo transdermal system
Outcomes	• Mean reduction of hot flushes per week from baseline • Adverse events based on review of participant diary cards
Notes	Data on effectiveness are skewed, reported in additional tables

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Authors reported that "each centre was packaged according to a computer-generated randomisation code"
Allocation concealment (selection bias)	Unclear risk	Not reported

Speroff 1996a (*Continued*)

Blinding of participants and personnel (performance bias) All outcomes	Low risk	Authors reported that: "E2 and placebo transdermal systems were identical in appearance, and investigators, study staff, and subjects remained blinded to treatment until the study was completed and the results were analysed"
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Authors reported that: "E2 and placebo transdermal systems were identical in appearance, and investigators, study staff, and subjects remained blinded to treatment until the study was completed and the results were analysed"
Incomplete outcome data (attrition bias) All outcomes	High risk	43/213 (20%) participants not included in analysis
Selective reporting (reporting bias)	Low risk	No protocol available, but reports all expected outcomes
Other bias	Unclear risk	Reported source of funding per Warner Chilcott

Speroff 1996b

Methods	The trial publication reports two independent RTCTs. Speroff 1996b is reported as Study 2 in the publication Multicentre randomised, double-blind, placebo-controlled trial Number of centres: 17 in USA Number of women randomised: 111 Number of women analysed: 91 Statistical analysis: analysis of covariance was used to compare the frequency of hot flushes per week between treatment groups within each study for each of the 12 weeks of the study Estradiol-estrone ratios of mean observed concentrations were calculated at each sampling time Duration of trial: 12 weeks Withdrawals and losses to follow-up: 20/111 (18%) Conflict of interest: this study was supported by Warner Chilcott, a division of Galen Holdings PLC, which has developed this product
Participants	Naturally menopausal women, all with prior hysterectomy Inclusion criteria: women with 56-140 hot flushes per week, had undergone hysterectomy, were at least 50 years old and naturally menopausal or at least 35 years old and surgically menopausal (bilateral oophorectomy at least 1 month before entry), had serum E2 concentrations of 20 pg/mL or less, and had FSH levels of 50 mlU/mL or more Exclusion criteria: women with contraindications to beta-estradiol replacement therapy or with a skin condition that might be exacerbated by the use of a transdermal system or that might mask a dermatologic reaction to the system

Speroff 1996b (*Continued*)

Interventions	Transdermal system delivering 0.02 mg and 0.04 mg of beta-estradiol versus placebo transdermal system
Outcomes	• Mean reduction of hot flushes per week from baseline • Adverse events based on review of participant diary cards
Notes	Data on effectiveness are skewed, reported in additional tables

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Authors reported that "each centre was packaged according to a computer-generated randomisation code"
Allocation concealment (selection bias)	Unclear risk	Not reported
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Authors reported that: "E2 and placebo transdermal systems were identical in appearance, and investigators, study staff, and subjects remained blinded to treatment until the study was completed and the results were analysed"
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Authors reported that: "E2 and placebo transdermal systems were identical in appearance, and investigators, study staff, and subjects remained blinded to treatment until the study was completed and the results were analysed"
Incomplete outcome data (attrition bias) All outcomes	High risk	20/111 (18%) participants not included in analysis. High rate of losses in the placebo group (30%) compared to the intervention group (7% and 11%)
Selective reporting (reporting bias)	Low risk	No protocol available, but reports all expected outcomes
Other bias	Unclear risk	No wash out period. Reported source of funding per Warner Chilcott

Utian 1999

Methods	Multicentre, double-blind, randomised, placebo-controlled trial, phase-III clinical trial in 4 parallel groups Number of centres: 26 in USA Number of women randomised: 196 Number of women analysed: 167 Statistical analysis: Homogeneity between the 4 treatment groups at baseline was tested statistically with $\chi 2$ tests, the Fisher exact test, Cochrane Mantel-Haenszel tests, or nonparametric analysis of variance. Within-group comparisons were performed with the Wilcoxon signed rank test. Between-group comparisons for vasomotor symptoms and laboratory tests were performed with analysis of variance, followed by comparisons against placebo for vasomotor symptoms by means of the Dunnett test. 2-sided tests with a significance level of 5% were used. Percentages were calculated on the basis of the number of participants on whom data were available. Mean values are given with the SD unless otherwise stated Withdrawals and losses to follow-up: n = 20. 7 because of adverse events Duration of trial: 12 weeks Conflict of interest: not stated
Participants	Women with signs of estrogen deficiency, particularly vasomotor symptoms consecutive to natural or surgical menopause Inclusion criteria: amenorrhoeic for at least 12 months, or 6 months with an estradiol level < 20 pg/mL and FSH level of 50 mIU/mL, or ovariectomised (bilateral oophorectomy for a benign reason) for at least 6 weeks. Particpants were to be highly symptomatic, with an average of at least 56 moderate to severe vasomotor symptoms per week (hot flushes and nocturnal sweating) during the 14-day symptom self-evaluation period that preceded the baseline visit Exclusion criteria: participants were not to have any marked benign mastopathy or any malignant growth, any major abnormality in the vaginal (Papanicolaou) smear, or any endometrial hyperplasia or carcinoma contraindicating estrogen replacement therapy. In addition, they were not to have any severe hepatic disorder, severe hyperlipidaemia, diabetes mellitus, or severe anaemia, as shown by the results of laboratory tests. Further exclusion criteria were known disease or allergy contraindicating estrogen therapy, chronic skin disease or history of cutaneous contact allergy, and cutaneous lesions on the buttocks
Interventions	Transdermal beta-estradiol patch 0.025 or 0.05 or 0.1 mg/d versus placebo
Outcomes	• Mean daily change from baseline of frequency and severity of moderate to severe hot flushes • Adverse events (including local skin reactions and genital bleeding). At each study visit, patch application sites were examined, information from the participant's diary was recorded, and participants were questioned about vasomotor symptoms, local skin reactions, patch detachments, genital bleeding, and adverse events
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement

Utian 1999 (*Continued*)

Random sequence generation (selection bias)	Low risk	Author reported "Study drug assignment was carried out by means of a centralized randomization system"
Allocation concealment (selection bias)	Unclear risk	No specific discussion on allocation concealment
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Author reported double-blind "Three different sizes of the placebo patches were used, matching each size of the active transdermal systems"
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	No specific discussion
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Author reported "Efficacy was analyzed for all treated patients evaluable for efficacy up to week 4 and, for patients, evaluable for efficacy up to week 12 or 13 as defined in the protocol. Similar conclusions were obtained, only the results for patients evaluable for efficacy up to week 12 or 13 are presented here..." Losses: 20/167 (12%), no ITT analysis
Selective reporting (reporting bias)	Low risk	No protocol available, but reports all expected outcomes
Other bias	Low risk	The study appears to be free of other sources of bias

von Holst 2000

Methods	Multicentre, double-blind, randomised, placebo-controlled trial Number of centres: 34 in Germany Number of women randomised; 186 Number of women analysed: 173 Statistical analysis: all statistical tests were performed 2-sided with a significance level at 5%. Differences between groups were quantified by calculating 2-tailed 95% CIs. Treatment groups were compared using the Wilcoxon rank-sum test. Additionally, post-menopausal complaints were analysed based on changes in specified symptoms after 3 treatment cycles during the double-blind, placebo-controlled phase II using the O'Brien method (non-parametric set-up) Withdrawals and losses to follow-up: n = not reported. 18 because of adverse events (7 placebo) Duration of trial: 12 weeks Conflict of interest: none declared
Participants	Hysterectomised women with postmenopausal symptoms Inclusion criteria: surgical postmenopausal status (bilateral oophorectomy) more than 3 months prior to inclusion in the study. Normal gynaecological history and examination, serum estradiol (E2) levels less than 30 pg/ml and FSH levels greater than 30 IU/ml. All of the women had at least 20 hot flushes per week within the last month prior to screening and an overall Kupperman Index greater than 20 Exclusion criteria: all those who had received any of the following prior to enrolment: sex hormones taken orally within the last 28 days; locally-applied sex hormones within the last 21 days or injectable sex hormones within the last 6 months
Interventions	Transdermal beta-estradiol 0.05 mg/d versus placebo
Outcomes	• Mean weekly change in frequency and severity of moderate to severe vasomotor symptoms from baseline • Adverse events whether reported by the participant or observed by the investigator were recorded
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Authors reported "...multicentre, randomized, double blind clinical study with an initial screening phase (phase I), a 3-month double-blind placebo-controlled phase (phase II) and a 3-month open follow-up phase (phase III).." No specific discussion
Allocation concealment (selection bias)	Unclear risk	Not reported

von Holst 2000 (*Continued*)

Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	No specific discussion
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	No specific discussion
Incomplete outcome data (attrition bias) All outcomes	Low risk	Losses: 18/173 (10%). ITT analysis: yes
Selective reporting (reporting bias)	Low risk	No protocol available, but reports all expected outcomes
Other bias	Low risk	Standard deviations not reported. These were imputed by the review authors

Wiklund 1993

Methods	Multicentre, double-blind, randomised, placebo-controlled trial Number of centres: 15 in Sweden Number of women randomised: 242 Number of women analysed: 223 Statistical analysis: the mean individual differences in quality of life and self-rated postmenopausal symptoms before and after therapy were tested with Wilcoxon's test. Only 2-tailed tests were used, and the underlying distributions were all close to the Gaussian one (without bimodal or skewed patterns). The first analysis was based on the overall or total scores. In cases of statistically significant differences $P < 0.01$) a second analysis was performed in a similar fashion with the sub scales or dimensions of the questionnaires Withdrawals and losses to follow-up: n = not reported. 18 because of adverse events Duration of trial: 12 weeks Conflict of interest: none declared
Participants	Postmenopausal women between 45 and 65 years old requiring HRT for climacteric symptoms were included Inclusion criteria: all the women had had their last menstruation at least 6 months previously Exclusion criteria: women with surgically induced menopause, previous or current estrogen-dependent tumour, other current malignant or life-threatening disease, severe metabolic, endocrine, or gastrointestinal disease, concomitant heart disease, insulin-treated diabetes, uncontrolled hypertension, endometriosis, undiagnosed vaginal bleeding, active skin disease, and unstable medical conditions such as rheumatoid arthritis or chronic obstructive lung disease. Women with psychiatric disorders and/or those receiving continuous tranquilliser or antidepressant therapy
Interventions	Transdermal beta-estradiol 0.05 mg/d versus placebo

Wiklund 1993 (*Continued*)

Outcomes	Mean change from baseline for vasomotor symptoms score (Index Kupperman) and quality-of-life measures during treatment measure used the questionnaires: Psychological General Well-Being Index, Nottingham Health Profile and Women's Health Questionnaire, McCoy Sex Scale. Self-rating scale of climacteric symptoms (VAS). Any events or side effects reported by the participant during the study were graded according to severity (mild, moderate, severe) and related to the trial treatment	
Notes		
Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Authors reported "The women were blindly and randomly allocated to either transdermal beta-estradiol therapy, 50 J.Lg/24 hours, or placebo given as patches twice a week..."
Allocation concealment (selection bias)	Unclear risk	Not reported
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Double-blind, matching placebo
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Described in detail
Incomplete outcome data (attrition bias) All outcomes	Low risk	Losses: 19/223 (8%)
Selective reporting (reporting bias)	Low risk	The study protocol is not available but it is clear that the published reports include all expected outcomes
Other bias	Low risk	The study appears to be free of other sources of bias

BMI: body mass index
FSH: follicle stimulating hormone
HT: hormone therapy
HRT: hormone replacement therapy
ITT: intention-to-treat
VAS: visual analogue scale

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Akhila 2006	Comparison is oral conjugated equine estrogen 0.625 mg versus gel beta-estradiol 0.75 mg versus patch beta-estradiol 1.8 mg plus medroxyprogesterone acetate 2,5 mg, which does not meet the selection criteria
Bachmann 2003	This study is not a RCT
Bachmann 2007	Comparison is beta-estradiol gel + levonorgestrel vs placebo, which does not meet the selection criteria for this review
Bachmann 2008	Comparison is beta-estradiol gel + levonorgestrel vs placebo, which does not meet the selection criteria of this review
Ben-Chetrit 2005	Comparison is ring beta-estradiol 0.36 g + progesterone 3.6 g versus ring beta-estradiol 0.36 g + progesterone 1.8 g , which does not meet the selection criteria of this review
Carranza-Lira 2006	Comparison is esterified conjugated estrogens (ECE) 0.156 mg/d (ultra low-dose) versus ECE 0.312 mg/d (low-dose), which does not meet the selection criteria of this review
Castelo-Branco 2010	This study is not a RCT
Chung 1996	Included participants without hot flushes
Conaway 2011	This study is not a RCT
Cortes-Bonilla 2015	Comparison is parenteral 0.5mg beta-estradiol+15mg progesterone versus 1mg parenteral beta-estradiol +20mg progesterone versus parenteral 1mg beta-estradiol +30mg progesterone,which does not meet the selection criteria of this review
Darj 1991	Compares treatment with different doses of micronised progesterone, all in combination with 0.2mg beta-estradiol plus. No group had placebo or non-bioidentical hormones
Diem 2006	Does not meet the selection criteria for this review as to the type of participants
Ettinger 2007	This study is not a RCT
Files 2011	This study is not a RCT
Formby 2011	This study is not a RCT
Ganz 2002	This study is not a RCT
Gass 2004	Comparison is beta-estradiol buccal tablets + medroxyprogesterone acetate vs placebo, which does not meet the selection criteria for this review
Hedrick 2009	Comparison 0.25, 0.5, or 1.0 g beta-estradiol 0.1% gel plus10 mg/d oral medroxyprogesterone for 14 days versus placebo,which does not meet the selection criteria for this review

(*Continued*)

Iftikhar 2011	This study is not a RCT
Jensen 1987	Participant group do not meet review inclusion criteria. Participants were part of a study of post-menopausal osteoporosis and did not all have vasomotor symptoms (incidence of hot flushes was 80% in one group and 93% in the other group at study commencement)
Lacut 2004	Participant group do not meet review inclusion criteria. Purpose of study was to investigate role of transdermal estrogen for preventing heart disease, and vasomotor symptoms were not an inclusion criterion; not stated whether women were symptomatic
Lindh 2004	Comparison is beta-estradiol + exercise vs sedentary women, which does not meet the selection criteria for this review
Lopes 2000	Comparison intranasal beta-estradiol 300 mg/day or patch 50 mg/day plus dydrogesterone 10 or 20 mg/d for 14 days per 28-day cycle, which does not meet the selection criteria for this review
Lubbert 1997	Comparison is beta-estradiol matrix patch (50 mg 17 /beta-estradiol/day) twice weekly, either continuously (8 patches/cycle) vs cyclically (6 patches/cycle, i.e. 3 weeks on, 1 week off) with or without an oral progestogen, which does not meet the inclusion criteria for this review
MacLennan 2009	This study is not a RCT
Marslew 1991	Comparison is 2 mg oestradiol valerate combined with cyproterone acetate, medroxyprogesterone acetate or levonorgestrel vs I.5 mg beta-oestradiol combined with desogestrel vs placebo, which does not meet the selection criteria for this review
Marslew 1994	Comparison is beta-oestradiol + cyproterone acetate vs beta-oestradiol + desogestrel which does not meet the selection criteria for this review
Mather 2000	Comparison is beta-estradiol vs beta-estradiol + progesterone vs progesterone which does not meet the selection criteria for this review. Outcomes do not include hot flushes
Mirkin 2015	This study is not an RCT. This report summarizes the methodology of the REPLENISH trial
Mizumuna 2011	Comparison is transdermal beta-estradiol 1.08 mg/d + medroxyprogesterone acetate or 0.9g /day transdermal beta-estradiol+ medroxyprogesterone acetate 5 mg vs placebo which does not meet the selection criteria for this review
Odabasi 2007	Comparison is intranasal beta-estradiol + progesterone gel vs beta-estradiol + progesterone gel, which does not meet the selection criteria for this review
Panay 2007	Comparison is beta-estradiol + norethisterone vs placebo which does not meet the selection criteria for this review
Pélissier 2001	Comparison is beta-estradiol patch + oral chlormadinone acetate versus beta- estradiol patch+ oral micronized progesterone so bioidentical versus bioidentical , which does not meet the selection criteria for this review

(*Continued*)

Rosano 2000	Comparison is beta-estradiol + progesterone gel vs medroxyprogesterone which does not meet the selection criteria for this review
Ryan 2001	Comparison is conjugated equine estrogens + medroxyprogesterone acetate vs conjugated equine + progesterone which does not meet the selection criteria for this review
Serfaty 2003	This study is not a RCT
Sitruk 2007	This study is not a RCT
Siyam 2013	This study is not a RCT
Skarsgard 2000	This study is not a RCT
Somunkiran 2007	Comparison is oral beta-estradiol vs tibolone which does not meet the selection criteria for this review
Sood 2011	This study is not a RCT
Sood 2013	Outcomes do not include hot flushes
Studd 1999	Comparison is intranasal beta-estradiol + medroxyprogesterone acetate 5 mg vs estradiol valerate 1 mg or 2 mg + medroxyprogesterone acetate 5 mg vs placebo which does not meet the selection criteria for s this review
Suvanto-luukkonen 1997	This study is not a RCT
Vartiainen 1993	This study is not a RCT
Veerus 2013	Comparison is 0.625 conjugated beta-estradiol + medroxyprogesterone acetate 2.5 mg or 0.625 conjugated beta-estradiol + medroxyprogesterone acetate 5 mg vs placebo which does not meet the selection criteria for s this review
Whelan 2013	Comparison is progesterone cream vs placebo which does meet the selection criteria for this review
Wihlbacka 2005	Comparison is beta-estradiol vs beta-estradiol + progesterone which does not meet the selection criteria for this review
Wolfe 1994	Comparison is Norgestrel dl or beta-estradiol or Norgestrel dl + beta-estradiol vs placebo which does not meet the selection criteria for this review
Yesildaglar 2004	Does not meet the selection criteria for this review as to the type of participants

DATA AND ANALYSES

Comparison 1. Transdermal beta estradiol patch vs placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Frequency of hot flushes	4	793	Std. Mean Difference (IV, Fixed, 95% CI)	-0.68 [-0.83, -0.53]
2 Intensity of hot flushes	2	393	Mean Difference (IV, Fixed, 95% CI)	-19.94 [-24.86, -15.02]
2.1 Beta-estradiol patch 0.025 mg/d	2	198	Mean Difference (IV, Fixed, 95% CI)	-17.49 [-24.61, -10.36]
2.2 Beta-estradiol patch 0.0375 mg/d	1	78	Mean Difference (IV, Fixed, 95% CI)	-11.0 [-22.59, 0.59]
2.3 Beta-estradiol patch 0.05 mg/d	1	117	Mean Difference (IV, Fixed, 95% CI)	-28.0 [-36.38, -19.62]
3 Adverse effects	9	1822	Odds Ratio (M-H, Random, 95% CI)	2.14 [1.29, 3.54]
3.1 Beta-estradiol patch 0.014 mg/d	1	160	Odds Ratio (M-H, Random, 95% CI)	1.85 [0.95, 3.59]
3.2 Beta-estradiol patch 0.025 mg/d	3	256	Odds Ratio (M-H, Random, 95% CI)	1.13 [0.33, 3.86]
3.3 Beta-estradiol patch 0.0375 mg/d	2	335	Odds Ratio (M-H, Random, 95% CI)	1.40 [0.88, 2.21]
3.4 Beta-estradiol patch 0.05 mg/d	6	845	Odds Ratio (M-H, Random, 95% CI)	1.94 [0.97, 3.87]
3.5 Beta-estradiol 0.10 mg/d	2	226	Odds Ratio (M-H, Random, 95% CI)	10.95 [5.62, 21.31]
4 Quality of life - beta-estradiol 0.014 mg/d	1	165	Mean Difference (IV, Fixed, 95% CI)	0.0 [-0.35, 0.35]
5 Quality of life - beta-estradiol 0.05 mg/d	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
5.1 Nottingham Health Profile	1	223	Mean Difference (IV, Fixed, 95% CI)	-40.90 [-61.82, -19.98]
5.2 Psychological general well-being	1	223	Mean Difference (IV, Fixed, 95% CI)	-7.0 [-11.90, -2.10]
5.3 Women's Health Questionnaire	1	223	Mean Difference (IV, Fixed, 95% CI)	-11.4 [-14.41, -8.39]
5.4 McCoy Sex Scale	1	223	Mean Difference (IV, Fixed, 95% CI)	-3.80 [-5.54, -2.06]
5.5 Self-rated symptoms	1	223	Mean Difference (IV, Fixed, 95% CI)	-157.9 [-200.37, -115.43]

Comparison 2. Transdermal beta-estradiol gel vs placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Adverse effects	3	1086	Odds Ratio (M-H, Fixed, 95% CI)	1.41 [1.09, 1.83]
1.1 Beta-estradiol gel 0.27 mg/d	1	190	Odds Ratio (M-H, Fixed, 95% CI)	1.55 [0.86, 2.80]
1.2 Beta-estradiol gel 0.37 mg/d	1	191	Odds Ratio (M-H, Fixed, 95% CI)	1.22 [0.67, 2.19]
1.3 Beta-estradiol gel 0.5 mg/d	1	182	Odds Ratio (M-H, Fixed, 95% CI)	1.10 [0.56, 2.16]
1.4 Beta-estradiol gel 0.75 mg/d	1	111	Odds Ratio (M-H, Fixed, 95% CI)	0.94 [0.41, 2.15]
1.5 Beta-estradiol gel 1.0 mg/d	1	188	Odds Ratio (M-H, Fixed, 95% CI)	1.42 [0.72, 2.79]
1.6 Beta-estradiol gel 1.5 mg/d	2	224	Odds Ratio (M-H, Fixed, 95% CI)	2.11 [1.20, 3.69]

Comparison 3. Oral beta-estradiol vs placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Frequency of hot flushes	2	356	Std. Mean Difference (IV, Fixed, 95% CI)	-0.80 [-1.03, -0.57]
1.1 Beta-estradiol oral 0.5 mg/d	2	178	Std. Mean Difference (IV, Fixed, 95% CI)	-0.62 [-0.94, -0.30]
1.2 Beta-estradiol oral 1.0 mg/d	2	178	Std. Mean Difference (IV, Fixed, 95% CI)	-0.98 [-1.31, -0.65]
2 Adverse effects	3	433	Odds Ratio (M-H, Fixed, 95% CI)	1.28 [0.84, 1.96]
2.1 Beta-estradiol oral 0.5 mg/d	2	178	Odds Ratio (M-H, Fixed, 95% CI)	1.00 [0.52, 1.93]
2.2 Beta-estradiol oral 1.0 mg/d	3	217	Odds Ratio (M-H, Fixed, 95% CI)	1.48 [0.81, 2.73]
2.3 Beta-estradiol oral 2.0 mg/d	1	38	Odds Ratio (M-H, Fixed, 95% CI)	1.88 [0.41, 8.63]

Comparison 4. Topical beta-estradiol emulsion vs placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Adverse effects	1	200	Odds Ratio (M-H, Fixed, 95% CI)	1.46 [0.80, 2.66]
1.1 Micellar nanoparticle beta-estradiol 8.6 mg/day	1	200	Odds Ratio (M-H, Fixed, 95% CI)	1.46 [0.80, 2.66]

Comparison 5. Intranasal beta-estradiol vs placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Frequency of hot flushes	1	458	Mean Difference (IV, Fixed, 95% CI)	-3.04 [-4.05, -2.03]
1.1 Beta-estradiol 0.021mg/d	1	154	Mean Difference (IV, Fixed, 95% CI)	-3.34 [-4.92, -1.76]
1.2 Beta-estradiol 0.029 mg/d	1	152	Mean Difference (IV, Fixed, 95% CI)	-2.47 [-4.45, -0.49]
1.3 Beta-estradiol 0.04 mg/d	1	152	Mean Difference (IV, Fixed, 95% CI)	-3.12 [-4.86, -1.38]
2 Adverse effects	1	458	Odds Ratio (M-H, Fixed, 95% CI)	1.96 [1.26, 3.03]
2.1 Beta estradiol 0.021mg/d	1	154	Odds Ratio (M-H, Fixed, 95% CI)	2.03 [0.92, 4.49]
2.2 Beta estradiol 0.029 mg/d	1	152	Odds Ratio (M-H, Fixed, 95% CI)	1.99 [0.98, 4.01]
2.3 Beta estradiol 0.04 mg/d	1	152	Odds Ratio (M-H, Fixed, 95% CI)	1.85 [0.85, 4.04]

Comparison 6. Oral beta-estradiol vs CEE 0.625

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Adverse effects	1	103	Odds Ratio (M-H, Fixed, 95% CI)	1.20 [0.50, 2.87]
1.1 Beta estradiol oral 1.0 mg/d	1	52	Odds Ratio (M-H, Fixed, 95% CI)	0.90 [0.25, 3.29]
1.2 Beta estradiol oral 2.0 mg/d	1	51	Odds Ratio (M-H, Fixed, 95% CI)	1.53 [0.46, 5.02]

ADDITIONAL TABLES

Table 1. Transdermal beta-estradiol patch versus placebo: frequency of hot flushes

Comparison	Measure	Transdermal beta-estradiol	Placebo	P value for difference between groups*

Table 1. Transdermal beta-estradiol patch versus placebo: frequency of hot flushes (*Continued*)

0.025 mg/d beta-estradiol transdermal patch vs placebo (De Aloysio 2000)	Mean reduction from baseline in number of hot flushes per day	83% (n = 47)	58% (n = 39)	P value < 0.05
0.0375 mg/d beta-estradiol transdermal patch vs placebo (De Aloysio 2000)		77% (n = 50)	58% (n = 39)	P value < 0.05
0.025 mg/d beta-estradiol transdermal patch vs placebo (Rovati 2000)		78% (n = 80)	50% (n = 80)	P value < 0.05.
0.05 mg/d beta-estradiol transdermal patch vs placebo (Rovati 2000)		93% (n = 77)	50% (n = 80)	P value < 0.05.
0.025 mg/d beta-estradiol transdermal patch vs placebo (Utian 1999)		86% (n = 42)	55% (n = 48)	P value < 0.05
0.05 mg/d beta-estradiol transdermal patch vs placebo (Utian 1999)		97% (n = 39)	55% (n = 48)	P value < 0.05
0.1 mg/d beta-estradiol transdermal patch vs placebo (Utian 1999)		97% (n = 39)	55% (n = 48)	P value < 0.05
0.02 mg/d beta-estradiol transdermal patch vs placebo (Speroff 1996a)	Mean number of hot flushes per week at end of study	13.4, SD 36.8 (n = 44)	21.5, SD 36.8 (n = 34)	P value = 0.088
0.04 mg/d beta-estradiol transdermal patch vs placebo (Speroff 1996a)		9.4, SD 35.1 (n = 39)	35.9, SD 35.1 (n = 33)	P value < 0.001
0.02 mg/d beta-estradiol transdermal patch vs placebo (Speroff 1996b)		13.4, SD 19.3 (n = 26)	29.8, SD 19.3-21.0 (n = 23)	P value = 0.004
0.04 mg/d beta estradiol transdermal patch vs placebo (Speroff 1996b)		12.4, SD 21.0 (n = 23)	29.8, SD 19.3-21.0 (n = 23)	P value = 0.006

Table 1. Transdermal beta-estradiol patch versus placebo: frequency of hot flushes (*Continued*)

0.014 mg/d beta-estradiol transdermal patch vs placebo (Haines 2009)	Mean reduction from baseline in number of hot flushes per month	55% (n = 80)	40% (n = 80)	P value < 0.01
0.05 mg/d beta estradiol transdermal patch vs placebo (Gordon 1995)	Reduction from baseline in number of hot flushes	n = 191 women included in analysis		P value < 0.05
0.10 mg/d beta-estradiol transdermal patch vs placebo (Gordon 1995)				P value < 0.05

*P value reported by authors
mg/d = milligrams per day

Table 2. Transdermal beta-estradiol patch versus placebo: hot flush intensity

Comparison	Transdermal estradiol	Placebo	P for difference between groups*
0.014 mg/d transdermal estradiol patch vs placebo (Haines 2009)	-5.8 (n = 80)	-8.4 (n = 80)	P value = 0.05

*P value reported by study authors
mg/d = milligrams per day

Table 3. Transdermal beta-estradiol gel versus placebo: frequency of hot flushes

Comparison	Measure	Transdermal estradiol	Placebo	P for difference between groups*
0.27 mg/d beta-estradiol transdermal gel vs placebo (Archer 2012)	Mean reduction from baseline in number of hot flushes per day	-5.86	-4.41	P value < 0.001
0.37 mg/d beta-estradiol transdermal gel vs placebo (Archer 2012)		-8.23	-4.41	P value < 0.001
0.75 mg/d beta-estradiol transdermal gel vs placebo (Archer 2003)		-7.8 (SD 4.17)	-5.7 (SD 4.19)	P value < 0.05
1.5 mg/d beta-estradiol transdermal gel vs placebo (Archer 2003)		-8.5 (SD 4.31)	-5.7 (SD 4.19)	P value < 0.05

Table 3. Transdermal beta-estradiol gel versus placebo: frequency of hot flushes (*Continued*)

0.52 mg/d beta-estradiol transdermal gel vs placebo (Simon 2007)	Proportion of women with relief of hot flushes at 12 weeks	61% (n = 132)	27% (n = 128)	P value < 0.001
0.52 mg/d beta-estradiol transdermal gel vs placebo (Simon 2007)		76% (n = 133)	27% (n = 128)	P value < 0.001
0.52 mg/d beta-estradiol transdermal gel vs placebo (Simon 2007)		77% (n = 64)	27% (n = 128)	P value < 0.001

*P value reported by authors
mg/d = milligrams per day

Table 4. Oral beta-estradiol versus placebo: frequency of hot flushes

Comparison (n)	Measure	Findings
1.0 mg/d oral beta-estradiol vs placebo (n = 27 vs 25) (Archer 1992)	Hot flush frequency at 12 weeks	91% reduction in intervention group vs 66% reduction in placebo group (P value < 0.01)
2.0 mg/d oral beta-estradiol vs placebo (n = 25 vs 25) (Archer 1992)		92% reduction in intervention group vs 66% reduction in placebo group (P value < 0.01)
0.25 mg/d oral 17 beta-estradiol vs placebo (n = 68 vs 66) (Notelovitz 2000a)	Hot flushes per week at week 12, reported in graphical form	All doses except 0.25 mg/d were significantly better than placebo at reducing moderate to severe hot flushes (P value < 0.001*), with a linear correlation for more benefit with higher dose
0.05 mg/d oral 17 beta-estradiol vs placebo (n = 64 vs 66) (Notelovitz 2000a)		
1.0 mg/d oral 17 beta-estradiol vs placebo (n = 67 vs 66) (Notelovitz 2000a)		
2.0 mg/d oral 17 beta-estradiol vs placebo (n = 68 vs 66) (Notelovitz 2000a)		

*P value reported by study authors
mg/d = milligrams per day

Table 5. Topical beta-estradiol emulsion versus placebo: frequency of hot flushes

Comparison	Measure	Topical estradiol emulsion	Placebo	P for difference between groups*
8.6 mg/d topical micellar nanoparticle beta-estradiol emulsion vs placebo (Simon 2006)	Hot flushes per day at 12 weeks	-11.1 (SD 6.8)	-7.2 (SD 4)	P value < 0.001

*P value reported by study authors
mg/d = milligrams per day

Table 6. Intranasal beta-estradiol versus placebo: quality-of-life score (Women's Health Questionnaire) change from baseline

Dimension	**150 µg/d (n = 114) mean (SD) P value***	**300 µg/d (n = 103) mean (SD) P value***	**Placebo (n = 118) mean (SD)**
Anxiety/depressed mood	-0.5 (12.6) ns	1.9 (11.8) ns	-1.6 (10.8)
Well-being	-1.0 (14.3) ns	3.4 (16.1) ns	-0.1 (14.8)
Somatic symptoms	0.8 (14.3) ns	2.0 (12.1) 0.012	-1.9 (14.8)
Memory/concentration	1.8 (16.2) 0.006	4.6 (17.4) < 0.001	-3.1 (16.9)
Vasomotor symptoms	25.1 (19.7) < 0.001	30.5 (29.6) < 0.001	2.3 (21.6)
Sleep problems	8.1 (21.2) < 0.001	8.2 (17.7) < 0.001	-1.9 (18.9)
Sexual behaviour	1.9 (21.0) 0.013	7.8 (20.3) < 0.001	-3.5 (17.0)
Menstrual symptoms	-2.3 (15.3) 0.003	-1.9 (16.1) 0.005	1.1 (0.0)

Data from Nielsen 2006
*P value compared with placebo, reported by study authors
ns = not statistically significant
mg/d = milligrams per day

Table 7. Transdermal beta -stradiol patch versus conjugated equine estrogens (CEE): frequency of hot flushes

Comparison	Measure	Beta-estradiol (n)	CEE (n)	P value for difference between groups*
0.05 mg/d transdermal beta-estradiol patch vs 0.0625 mg/d CEE (Good 1999)	Mean change from baseline in number of hot flushes per week, at 12 weeks Percentages in this table are estimated, as data presented in graphical form	87%	91%	P value > 0.05
0.10 mg/d transdermal beta-estradiol patch vs 0.0625 mg/d CEE (Good 1999)		90%	91%	P value > 0.05
0.05 mg/d transdermal beta-estradiol patch vs 1.25 mg/d CEE (Good 1999)		87%	93%	P value > 0.05
0.10 mg/d transdermal beta-estradiol patch vs 1.25 mg/d CEE (Good 1999)		90%	93%	P value > 0.05
0.05 mg/d transdermal beta-estradiol patch vs 0.0625 mg/d CEE (Gordon 1995)	Mean change from baseline in number of hot flushes per week, at 11 weeks	62.8% (n = 130)	67.3% (n = 136)	P value > 0.05
0.05 mg/d transdermal beta-estradiol patch vs 0.0625 mg/d CEE (Gordon 1995)		78.1% (n = 124)	67.3% (n = 136)	P value > 0.05

P value compared with placebo, reported by study authors
mg/d = milligrams per day

Table 8. Oral beta-estradiol versus conjugated equine estrogens (CEE): frequency of hot flushes

Comparison	Measure	Beta-estradiol (n)	CEE (n)	P value for difference between groups*
1.0 mg/d oral beta-estradiol vs 0.0625 mg/d CEE (Archer 1992)	Mean change from baseline in frequency of vasomotor symptoms, at 12 weeks	91% decrease from baseline (n = 27)	80% decrease from baseline (n = 25)	P value > 0.05
2.0 mg/d oral beta-estradiol vs 0.0625 mg/d				

Table 8. Oral beta-estradiol versus conjugated equine estrogens (CEE): frequency of hot flushes (*Continued*)

CEE (Archer 1992)				
1.0 mg/d oral beta-estradiol vs 1.25 mg/d CEE (Archer 1992)		91% decrease from baseline (n = 27)	95% decrease from baseline (n = 26)	P value > 0.05
2.0 mg/d oral beta-estradiol vs 1.25 mg/d CEE (Archer 1992)		92% decrease from baseline (n = 25)	95% decrease from baseline (n = 26)	P value > 0.05

P value compared with placebo, reported by study authors
mg/d = milligrams per day

CONTRIBUTIONS OF AUTHORS

Drafting the protocol: Ana Marcia IS Gaudard, Edina MK da Silva, Maria R. Torloni and Sulani Silva de Souza

Drafting the review: Ana Marcia IS Gaudard, Edina MK da Silva, Maria R. Torloni, Sulani Silva de Souza and Jane Marjoribanks

Development of search strategy: Ana Marcia IS Gaudard

Search for trials: Ana Marcia IS Gaudard

Selection of which trials to include: Ana Marcia IS Gaudard, Edina MK da Silva, Sulani Silva de Souza, Maria R Torloni

Extraction of data from trials: Ana Marcia IS Gaudard, Edina MK da Silva, Sulani Silva de Souza, Maria R Torloni and Jane Marjoribanks

Assessment of risk of bias in included studies: Ana Marcia IS Gaudard, Edina MK da Silva, Sulani Silva de Souza

Entry of data into RevMan: Ana Marcia IS Gaudard

DECLARATIONS OF INTEREST

Ana Marcia IS Gaudard: none known

Sulani Silva de Souza: none known

Maria ES Puga: none known

Maria R Torloni: none known

Jane Marjoribanks: none known

Edina MK da Silva: none known

SOURCES OF SUPPORT

Internal sources

- Brazilian Cochrane Centre, Brazil.

Brazilian Cochrane Centre to conduct the searches in LILACS

External sources

- None, Other.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

- The analyses were stratified by dose of BHT.
- We added a new sensitivity analysis.
- We decided not to do a sensitivity analysis by outlier status, as data-driven decisions are associated with high risk of bias.
- We noted explicitly in the Methods section that we conducted separate comparisons according to the route of administration of the intervention.

INDEX TERMS

Medical Subject Headings (MeSH)

Estradiol [adverse effects; *therapeutic use]; Estrogens [adverse effects; *therapeutic use]; Estrogens, Conjugated (USP) [adverse effects; *therapeutic use]; Hot Flashes [*drug therapy]; Randomized Controlled Trials as Topic

MeSH check words

Female; Humans